Springer Series in Computational Mathematics 51

Volker John

Finite Element Methods for Incompressible Flow Problems

Springer

Springer Series in Computational Mathematics

Volume 51

More information about this series at http://www.springer.com/series/797

Volker John

Finite Element Methods for Incompressible Flow Problems

Springer

Volker John
Weierstrass Institute for Applied Analysis
and Stochastics
Berlin, Germany

Fachbereich Mathematik und Informatik
Freie Universität Berlin
Berlin, Germany

ISSN 0179-3632 ISSN 2198-3712 (electronic)
Springer Series in Computational Mathematics
ISBN 978-3-319-45749-9 ISBN 978-3-319-45750-5 (eBook)
DOI 10.1007/978-3-319-45750-5

Library of Congress Control Number: 2016956572

Mathematics Subject Classification (2010): 65M60, 65N30, 35Q30, 76F65, 76D05, 76D07

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

For Anja and Josephine

Preface

Incompressible flow problems appear in many models of physical processes and applications. Their numerical simulation requires in particular a spatial discretization. Finite element methods belong to the mathematically best understood discretization techniques.

This monograph is devoted mainly to the mathematical aspects of finite element methods for incompressible flow problems. It addresses researchers, Ph.D. students, and even students aiming for the master's degree. The presentation of the material, in particular of the mathematical arguments, is performed in detail. This style was chosen in the hope to facilitate the understanding of the topic, especially for nonexperienced readers.

Most parts of this monograph were presented in three consecutive master's level courses taught at the Free University of Berlin, and this monograph is based on the corresponding lecture notes. First of all, I like to thank the students who attended these courses. Many of them wrote finally their master's thesis under my supervision. Then, I like to thank two collaborators of mine, Julia Novo (Madrid) and Gabriel R. Barrenechea (Glasgow), who read parts of this monograph and gave valuable suggestions for improvement. Above all, I like to thank my beloved wife Anja and my daughter Josephine for their continual encouragement. Their efforts to manage our daily life and to save me working time were an invaluable contribution for writing this monograph in the past 3 years.

Colbitz, Germany Volker John
July 2016

Contents

Chapter 1
Introduction

The behavior of incompressible fluids is modeled with the system of the incompressible Navier–Stokes equations. These equations describe the conservation of linear momentum and the conservation of mass. In the special case of a steady-state situation and large viscosity of the fluid, the Navier–Stokes equations can be approximated by the Stokes equations. Incompressible flow problems are not only of interest by themselves, but they are part of many complex models for describing phenomena in nature or processes in engineering and industry.

Usually it is not possible to find an analytic solution of the Stokes or Navier–Stokes equations such that numerical methods have to be employed for approximating the solution. To this end, a so-called discretization has to be applied to the equations, in the general case a temporal and a spatial discretization. Concerning the spatial discretization, this monograph considers finite element methods. Finite element methods are very popular and they are understood quite well from the mathematical point of view.

First applications of finite element methods for the simulation of the Stokes and Navier–Stokes equations were performed in the 1970s. Also the finite element analysis for these equations started in this decade, e.g., by introducing in Babuška (1971) and Brezzi (1974) the inf-sup condition which is a basis of the well-posedness of the continuous as well as of the finite element problem. The early works on the finite element analysis cumulated in the monograph (Girault and Raviart 1979). The extended version of this monograph, Girault and Raviart (1986), became the classical reference work. Three decades have been passed since its publication. Of course, in this time, there were many new developments and new results have been obtained. More recent monographs that study finite element methods for incompressible flow problems, or important aspects of this topic, include Layton (2008), Boffi et al. (2008), Elman et al. (2014).

This monograph covers on the one hand a wide scope, from the derivation of the Navier–Stokes equations to the simulation of turbulent flows. On the other hand, there are many topics whose detailed presentation would amount in a monograph

V. John, *Finite Element Methods for Incompressible Flow Problems*, Springer Series in Computational Mathematics 51, DOI 10.1007/978-3-319-45750-5_1

itself and they are only sketched here. The main emphasis of the current monograph is on mathematical issues. Besides many results for finite element methods, also a few fundamental results concerning the continuous equations are presented in detail, since a basic understanding of the analysis of the continuous problem provides a better understanding of the considered problem in its entirety.

A main feature of this monograph is the detailed presentation of the mathematical tools and of most of the proofs. This feature arose from the experience in sometimes spending (wasting) a lot of time in understanding certain steps in proofs that are written in the short form which is usual in the literature. Often, such steps would have been easy to understand if there would have been just an additional hint or one more line in the estimate. Thus, the presentation is mostly self-contained in the way that no other literature has to be used for understanding the majority of the mathematical results. Altogether, the monograph is directed to a broad audience: experienced researchers on this topic, young researchers, and master students. The latter point was successfully verified. Most parts of this monograph were presented in master courses held at the Free University of Berlin, in particular from 2013–2015. As a result, several master's theses were written on topics related to these courses.

1.1 Contents of this Monograph

Chapter 2 sketches the derivation of the incompressible Navier–Stokes equations on the basis of the conservation of mass and the conservation of linear momentum. Important properties of the stress tensor are derived from physical considerations. The non-dimensionalized equations are introduced and appropriate boundary conditions are discussed.

The following structure of this monograph is based on the inherent difficulties of the incompressible Navier–Stokes equations pointed out in Chap. 2.

- First, the coupling of velocity and pressure is studied:
  - Chapter 3 presents an abstract theory and discusses the choice of appropriate finite element spaces.
  - Chapter 4 applies the abstract theory to the Stokes equations.
- Second, the issue of dominant convection is also taken into account:
  - Chapter 5 studies this topic for the Oseen equations, which are a kind of linearized Navier–Stokes equations.
- Third, the nonlinearity of the Navier–Stokes equations is considered in addition to the other two difficulties:
  - Chapter 6 studies stationary flows that occur only for small Reynolds numbers.
  - Chapter 7 considers laminar flows that arise for medium Reynolds numbers.
  - Chapter 8 studies turbulent flows that occur for large Reynolds numbers.

The coupling of velocity and pressure in incompressible flow problems does not allow the straightforward use of arbitrary pairs of finite element spaces. For obtaining a well-posed problem, the spaces have to satisfy the so-called discrete inf-sup condition. This condition is derived in Chap. 3. The derivation is based on the study of the well-posedness of an abstract linear saddle point problem. The abstract theory is applied first to a continuous linear incompressible flow problem, thereby identifying appropriate function spaces for velocity and pressure. These spaces satisfy the so-called inf-sup condition. Then, it is discussed that the satisfaction of the inf-sup condition does not automatically lead to the satisfaction of the discrete inf-sup condition. Examples of velocity and pressure finite element spaces that do not satisfy this condition are given. Next, some techniques for proving the discrete inf-sup condition are presented and important inf-sup stable pairs of finite element spaces are introduced. For some pairs, the proof of the discrete inf-sup condition is presented. In addition, a way for computing the discrete inf-sup constant is described. The final section of this chapter discusses the Helmholtz decomposition.

Chapter 4 applies the theory developed in the previous chapter to the Stokes equations. The Stokes equations, being a system of linear equations, are the simplest model of incompressible flows, modeling only the flow caused by viscous forces. First, the existence, uniqueness, and stability of a weak solution is discussed. The next section presents results from the finite element error analysis. Conforming finite element methods are considered in the first part of this section and a low order non-conforming finite element discretization is studied in the second part. Some remarks concerning the implementation of the finite element methods are given. Next, a basic introduction to a posteriori error estimation is presented and its application for adaptive mesh refinement is sketched. It follows a presentation of methods that allow to circumvent the discrete inf-sup condition. Such methods enable the usage of the same finite element spaces with respect to the piecewise polynomials for velocity and pressure, which is appealing from the practical point of view. A detailed numerical analysis of one of these methods, the PSPG method, is provided and a couple of other methods are discussed briefly. Finite element methods satisfy in general the conservation of mass only approximately. This chapter concludes with a survey of methods that reduce the violation of mass conservation or even guarantee its conservation.

The Oseen equations, i.e., a linear equation with viscous (second order), convective (first order), and reactive (zeroth order) term are the topic of Chap. 5. These equations arise in various numerical methods for solving the Navier–Stokes equations. Usually, the convective or the reactive term dominate the viscous term. A major issue in the analysis consists in tracking the dependency of the stability and error bounds on the coefficients of the problem. After having established the existence and uniqueness of a solution of the equations, the standard Galerkin finite element method is studied. It turns out that the stability and error bounds become large if convection or reaction dominates. Numerical studies support this statement. For improving the numerical solutions, stabilized methods have to be applied. The analysis of a residual-based stabilized method, the SUPG/PSPG/grad-div method, is presented in detail and some further stabilized methods are reviewed briefly.

In Chap. 6, the first nonlinear model of an incompressible flow problem is studied—the steady-state Navier–Stokes equations. At the beginning of this chapter, the nonlinear term is investigated. Different forms of this term are introduced and various properties are derived. Then, the solution of the continuous steady-state Navier–Stokes equations is studied. It turns out that a unique solution can be expected only for sufficiently small external forces, which do not depend on time, and sufficiently large viscosity. For this situation, a finite element error analysis is presented, with the emphasis on bounding the nonlinear term. Next, iterative methods for solving the nonlinear problem are discussed. The final section of this chapter presents the Dual Weighted Residual (DWR) method. This method is an approach for the a posteriori error estimation with respect to quantities of interest.

Chapter 7 starts with the investigation of the time-dependent incompressible Navier–Stokes equations. From the point of view of finite element discretizations, so-called laminar flows are considered, i.e., flows where a standard Galerkin finite element method is applicable. At the beginning of this chapter, a short introduction into the analysis concerning the existence and uniqueness of a weak solution of the time-dependent incompressible Navier–Stokes equations is given. In particular, the mathematical reason is highlighted that prevents to prove the uniqueness in the practically relevant three-dimensional case. Then, the finite element error analysis for the Galerkin finite element method in the so-called continuous-in-time case is presented, i.e., without the consideration of a discretization with respect to time. For practical simulations, a temporal discretization has to be applied. The next part of this chapter introduces a number of time stepping schemes that require the solution of a coupled velocity-pressure problem in each discrete time. In particular, θ-schemes are discussed in detail. It follows the presentation of a finite element error analysis for the fully discretized equations at the example of the backward Euler scheme. The approaches presented so far in this chapter require the solution of saddle point problems, which might be computational expensive. Projection methods, which circumvent the solution of such problems, are presented in the last section of this chapter. In these methods, only scalar equations for each component of the velocity field and for the pressure have to be solved.

The topic of Chap. 8 is the simulation of turbulent flows. There is no mathematical definition of what is a turbulent flow. Thus, this chapter starts with a description of characteristics of flow fields that are considered to be turbulent. In addition, a mathematical approach for describing turbulence is sketched. It turns out that turbulent flows possess scales that are much too small to be representable on grids with affordable fineness. The impact of these scales on the resolvable scales has to be modeled with a so-called turbulence model. The bulk of this chapter presents turbulence models that allow mathematical or numerical analysis or whose derivation is based on mathematical arguments. A very popular approach for turbulence modeling is large eddy simulation (LES). LES aims at simulating only large (resolved) scales that are defined by spatial averaging. In the first section on LES, the derivation of equations for these scales is discussed, in particular the underlying assumption of commuting differentiation and spatial averaging. It is shown that usually commutation errors occur that are not negligible. The next section presents the most popular LES model, the Smagorinsky model. For the Smagorinsky model, a well developed mathematical and numerical analysis is available. Then, LES models are described that are derived on the basis of approximations in wave number space. The final section on LES considers so-called Approximate Deconvolution models (ADM). As next turbulence model, the Leray-α model is presented. This model is based on a regularization of the velocity field. Afterward, the Navier–Stokes-α model is considered. Its derivation is performed by studying a Lagrangian functional and the corresponding trajectory. The last class of turbulence models that is discussed is the class of Variational Multiscale (VMS) methods. VMS methods define the large scales, which should be simulated, in a different way than LES models, namely by projections in appropriate function spaces. Two principal types of VMS methods can be distinguished, those based on a two-scale decomposition and those using a three-scale decomposition of the flow field. Five different realizations of VMS methods are described in detail. The final section of Chap. 8 presents a few numerical studies of turbulent flow simulations using the Smagorinsky model and one representative of the VMS models.

The linearization and discretization of the incompressible Navier–Stokes equations results for many methods in coupled algebraic systems for velocity and pressure. Chapter 9 gives a brief introduction into solvers for such equations. One can distinguish between sparse direct solvers and iterative solvers, where the latter solvers have to be used with appropriate preconditioners. Some emphasis in the presentation is on the preconditioner that was used for simulating most of the numerical examples presented in this monograph, namely a coupled multigrid method.

Appendix A provides some basic notations from functional analysis. A number of inequalities and theorems are given that are used in the analysis and numerical analysis presented in this monograph. Some basics of the finite element method are provided in Appendix B. In particular, those finite element spaces are described in some detail that are used for discretizing incompressible flow problems. The approximation of functions from Sobolev spaces with finite element functions by interpolation or projection is the topic of Appendix C. The corresponding estimates are heavily used in the finite element error analysis. Finally, Appendix D describes a number of examples for numerical simulations, which are divided into three groups:

- examples for steady-state flow problems,
- examples for laminar time-dependent flow problems,
- examples for turbulent flow problems.

The described examples were utilized for performing numerical simulations whose results are presented in this monograph.

The master courses held at the Free University of Berlin covered the following topics:

- Course 1: Chaps. 2, and 3, Sect. 4.1–4.3,
- Course 2: Sects. 4.4–4.6, Chaps. 5–7, and 9,
- Course 3: Chap. 8.

Of course, the presentation in these courses concentrated on the most important issues of each topic.

Chapter 2
The Navier–Stokes Equations as Model for Incompressible Flows

Remark 2.1 (Basic Principles and Variables) The basic equations of fluid dynamics are called Navier–Stokes equations. In the case of an isothermal flow, i.e., a flow at constant temperature, they represent two physical conservation laws: the conservation of mass and the conservation of linear momentum. There are various ways for deriving these equations. Here, the classical one of continuum mechanics will be outlined. This approach contains some heuristic steps.

The flow will be described with the variables

- $\rho(t, \boldsymbol{x})$: density [$^{\mathrm{kg}}/_{\mathrm{m}^3}$],
- $\boldsymbol{v}(t, \boldsymbol{x})$: velocity [$^{\mathrm{m}}/_{\mathrm{s}}$],
- $P(t, \boldsymbol{x})$: pressure [$\mathrm{Pa} = {}^{\mathrm{N}}/_{\mathrm{m}^2}$],

which are assumed to be sufficiently smooth functions in the time interval $[0, T]$ and the domain $\Omega \subset \mathbb{R}^3$. □

2.1 The Conservation of Mass

Remark 2.2 (General Conservation Law) Let ω be an arbitrary open volume in Ω with sufficiently smooth surface $\partial\omega$, which is constant in time, and with mass

$$m(t) = \int_\omega \rho(t, \boldsymbol{x})\, d\boldsymbol{x}\ [\mathrm{kg}].$$

If mass in ω is conserved, the rate of change of mass in ω must be equal to the flux of mass $\rho\boldsymbol{v}(t, \boldsymbol{x})$ $[\mathrm{kg}/(\mathrm{m}^2\mathrm{s})]$ across the boundary $\partial\omega$ of ω

$$\frac{d}{dt} m(t) = \frac{d}{dt} \int_\omega \rho(t, \boldsymbol{x})\, d\boldsymbol{x} = -\int_{\partial\omega} (\rho\boldsymbol{v})\,(t, \boldsymbol{s}) \cdot \boldsymbol{n}(\boldsymbol{s})\, ds, \tag{2.1}$$

V. John, *Finite Element Methods for Incompressible Flow Problems*, Springer Series in Computational Mathematics 51, DOI 10.1007/978-3-319-45750-5_2

where $\boldsymbol{n}(\boldsymbol{s})$ is the outward pointing unit normal on $\boldsymbol{s} \in \partial\omega$. Since all functions and $\partial\omega$ are assumed to be sufficiently smooth, the divergence theorem can be applied (integration by parts), which gives

$$\int_{\omega} \nabla \cdot (\rho \boldsymbol{v})\,(t, \boldsymbol{x})\; d\boldsymbol{x} = \int_{\partial\omega} (\rho \boldsymbol{v})\,(t, \boldsymbol{s}) \cdot \boldsymbol{n}(\boldsymbol{s})\; d\boldsymbol{s}.$$

Inserting this identity in (2.1) and changing differentiation with respect to time and integration with respect to space leads to

$$\int_{\omega} \left(\partial_t \rho(t, \boldsymbol{x}) + \nabla \cdot (\rho \boldsymbol{v})\,(t, \boldsymbol{x})\right)\; d\boldsymbol{x} = 0.$$

Since ω is an arbitrary volume, it follows that

$$(\partial_t \rho + \nabla \cdot (\rho \boldsymbol{v}))\,(t, \boldsymbol{x}) = 0\, forallst \in (0, T],\ \boldsymbol{x} \in \Omega. \tag{2.2}$$

This relation is the first equation of fluid dynamics, which is called continuity equation. □

Remark 2.3 (Time-Dependent Domain) It is also possible to consider a time-dependent domain $\omega(t)$. In this case, the Reynolds transport theorem can be applied. Let $\phi(t, \boldsymbol{x})$ be a sufficiently smooth function defined on an arbitrary volume $\omega(t)$ with sufficiently smooth boundary $\partial\omega(t)$, then the Reynolds transport theorem has the form

$$\frac{d}{dt} \int_{\omega(t)} \phi(t, \boldsymbol{x})\; d\boldsymbol{x} = \int_{\omega(t)} \partial_t \phi(t, \boldsymbol{x})\; d\boldsymbol{x} + \int_{\partial\omega(t)} (\phi \boldsymbol{v} \cdot \boldsymbol{n})\,(t, \boldsymbol{s})\; ds. \tag{2.3}$$

In the special case that $\phi(t, \boldsymbol{x})$ is the density, one gets for the change of mass

$$\frac{d}{dt} \int_{\omega(t)} \rho(t, \boldsymbol{x})\; d\boldsymbol{x} = \int_{\omega(t)} \partial_t \rho(t, \boldsymbol{x})\; d\boldsymbol{x} + \int_{\partial\omega(t)} (\rho \boldsymbol{v} \cdot \boldsymbol{n})\,(t, \boldsymbol{s})\; ds.$$

Conservation of mass and the divergence theorem yields

$$0 = \int_{\omega(t)} (\partial_t \rho + \nabla \cdot (\rho \boldsymbol{v}))\,(t, \boldsymbol{x})\; d\boldsymbol{x}.$$

Since $\omega(t)$ is assumed to be arbitrary, Eq. (2.2) follows. □

Remark 2.4 (Incompressible, Homogeneous Fluids) If the fluid is incompressible and homogeneous, i.e., composed of one fluid only, then $\rho(t, \boldsymbol{x}) = \rho > 0$ and (2.2) reduces to

$$\left(\partial_x v_1 + \partial_y v_2 + \partial_z v_3\right)(t, \boldsymbol{x}) = \nabla \cdot \boldsymbol{v}(t, \boldsymbol{x}) = 0\, forallst \in (0, T],\ \boldsymbol{x} \in \Omega, \tag{2.4}$$

where

$$\boldsymbol{v}(t,\boldsymbol{x}) = \begin{pmatrix} v_1(t,\boldsymbol{x}) \\ v_2(t,\boldsymbol{x}) \\ v_3(t,\boldsymbol{x}) \end{pmatrix}.$$

Thus, the conservation of mass for an incompressible, homogeneous fluid imposes a constraint on the velocity only. □

2.2 The Conservation of Linear Momentum

Remark 2.5 (Newton's Second Law of Motion) The conservation of linear momentum is the formulation of Newton's second law of motion

$$\text{net force} = \text{mass} \times \text{acceleration} \tag{2.5}$$

for flows. It states that the rate of change of the linear momentum must be equal to the net force acting on a collection of fluid particles. □

Remark 2.6 (Conservation of Linear Momentum) The linear momentum in an arbitrary volume ω is given by

$$\int_\omega \rho \boldsymbol{v}(t,\boldsymbol{x})\, d\boldsymbol{x} \quad [\text{Ns}].$$

Then, the conservation of linear momentum in ω can be formulated analogously to the conservation of mass in (2.1)

$$\frac{d}{dt}\int_\omega \rho \boldsymbol{v}(t,\boldsymbol{x})\, d\boldsymbol{x} = -\int_{\partial\omega} (\rho\boldsymbol{v})\,(\boldsymbol{v}\cdot\boldsymbol{n})\,(t,\boldsymbol{s})\, d\boldsymbol{s} + \int_\omega \boldsymbol{f}_{\text{net}}(t,\boldsymbol{x})\, d\boldsymbol{x}\ [\text{N}],$$

where the term on the left-hand side describes the change of the momentum in ω, the first term on the right-hand side models the flux of momentum across the boundary of ω and $\boldsymbol{f}_{\text{net}}$ [N/m^3] represents the force density in ω. It is

$$\boldsymbol{v}(\boldsymbol{v}\cdot\boldsymbol{n}) = \begin{pmatrix} v_1v_1n_1 + v_1v_2n_2 + v_1v_3n_3 \\ v_2v_1n_1 + v_2v_2n_2 + v_2v_3n_3 \\ v_3v_1n_1 + v_3v_2n_2 + v_3v_3n_3 \end{pmatrix} = \boldsymbol{v}\boldsymbol{v}^T\boldsymbol{n}.$$

Applying integration by parts and changing differentiation with respect to time and integration on ω gives

$$\int_\omega \left(\partial_t(\rho\boldsymbol{v}) + \nabla\cdot\left(\rho\boldsymbol{v}\boldsymbol{v}^T\right)\right)(t,\boldsymbol{x})\, d\boldsymbol{x} = \int_\omega \boldsymbol{f}_{\text{net}}(t,\boldsymbol{x})\, d\boldsymbol{x}.$$

The product rule yields

$$\int_\omega \big(\partial_t \rho \boldsymbol{v} + \rho \partial_t \boldsymbol{v} \; + \; \boldsymbol{v}\boldsymbol{v}^T \nabla \rho + \rho (\nabla \cdot \boldsymbol{v})\boldsymbol{v} + \rho (\boldsymbol{v} \cdot \nabla)\boldsymbol{v}\big) (t, \boldsymbol{x}) \, d\boldsymbol{x}$$
$$= \int_\omega \boldsymbol{f}_{\text{net}}(t, \boldsymbol{x}) \, d\boldsymbol{x}. \tag{2.6}$$

In the usual notation $(\boldsymbol{v} \cdot \nabla)\boldsymbol{v}$, one can think of $\boldsymbol{v} \cdot \nabla = v_1 \partial_x + v_2 \partial_y + v_3 \partial_z$ acting on each component of $\boldsymbol{v}$. In the literature, one often finds the notation $\boldsymbol{v} \cdot \nabla \boldsymbol{v}$.

In the case of incompressible flows, i.e., ρ is constant, it is known that $\nabla \cdot \boldsymbol{v} = 0$, see (2.4), such that (2.6) simplifies to

$$\int_\omega \rho \left(\partial_t \boldsymbol{v} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v}\right)(t, \boldsymbol{x}) \, d\boldsymbol{x} = \int_\omega \boldsymbol{f}_{\text{net}}(t, \boldsymbol{x}) \, d\boldsymbol{x}.$$

Since ω was chosen to be arbitrary, one gets the conservation law

$$\rho \left(\partial_t \boldsymbol{v} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v}\right) = \boldsymbol{f}_{\text{net}} \qquad \forall \, t \in (0, T], \; \boldsymbol{x} \in \Omega.$$

The same conservation law can be derived for a time-dependent volume $\omega(t)$ using the Reynolds transport theorem (2.3). □

Remark 2.7 (External Forces) The forces acting on ω are composed of external (body) forces and internal forces. External forces include, e.g., gravity, buoyancy, and electromagnetic forces (in liquid metals). These forces are collected in a body force term

$$\int_\omega \boldsymbol{f}_{\text{ext}}(t, \boldsymbol{x}) \, d\boldsymbol{x}.$$

□

Remark 2.8 (Internal Forces, Cauchy's Principle, and the Stress Tensor) Internal forces are forces which a fluid exerts on itself. These forces include the pressure and the viscous drag that a 'fluid element' exerts on the adjacent element. The internal forces of a fluid are contact forces, i.e., they act on the surface of the fluid element ω. Let $\boldsymbol{t}$ $[\mathrm{N}/\mathrm{m}^2]$ denote this internal force vector, which is called Cauchy stress vector or torsion vector, then the contribution of the internal forces on ω is

$$\int_{\partial\omega} \boldsymbol{t}(t, \boldsymbol{s}) \, d\boldsymbol{s}.$$

Adding the external and internal forces, the equation for the conservation of linear momentum is, for an arbitrary constant-in-time volume ω,

$$\int_\omega \rho(t, \boldsymbol{x}) \left(\partial_t \boldsymbol{v} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v}\right)(t, \boldsymbol{x}) \, d\boldsymbol{x} = \int_\omega \boldsymbol{f}_{\text{ext}}(t, \boldsymbol{x}) \, d\boldsymbol{x} + \int_{\partial\omega} \boldsymbol{t}(t, \boldsymbol{s}) \, d\boldsymbol{s}. \tag{2.7}$$

The right-hand side of (2.7) describes the net force acting on and inside ω. Now, a detailed description of the internal forces represented by $\boldsymbol{t}(t, \boldsymbol{s})$ is necessary.

The foundation of continuum mechanics is the stress principle of Cauchy. The idea of Cauchy on internal contact forces was that on any (imaginary) plane on $\partial\omega$ there is a force that depends (geometrically) only on the orientation of the plane. Thus, it is $\boldsymbol{t} = \boldsymbol{t}(\boldsymbol{n})$, where $\boldsymbol{n}$ is the outward pointing unit normal vector of the imaginary plane.

Next, it will be discussed that $\boldsymbol{t}$ depends linearly on $\boldsymbol{n}$. To this end, consider a tetrahedron ω with the vertices $\boldsymbol{p}_0 = (0,0,0)^T, \boldsymbol{p}_1 = (x_1,0,0)^T, \boldsymbol{p}_2 = (0,y_2,0)^T$, $\boldsymbol{p}_3 = (0,0,z_3)^T$, and with $x_1, y_2, z_3 > 0$, see Fig. 2.1 for an illustration. Denote the plane containing $\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3$ by $\partial\omega^{(\boldsymbol{n})}$. The unit outward pointing normal of $\partial\omega^{(\boldsymbol{n})}$ is given by

$$\begin{aligned}\boldsymbol{n} &= \frac{(\boldsymbol{p}_2 - \boldsymbol{p}_1) \times (\boldsymbol{p}_3 - \boldsymbol{p}_1)}{\|(\boldsymbol{p}_2 - \boldsymbol{p}_1) \times (\boldsymbol{p}_3 - \boldsymbol{p}_1)\|_2} \\ &= \frac{1}{\|(\boldsymbol{p}_2 - \boldsymbol{p}_1) \times (\boldsymbol{p}_3 - \boldsymbol{p}_1)\|_2} \begin{pmatrix} y_2 z_3 \\ z_3 x_1 \\ x_1 y_2 \end{pmatrix} = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}. \end{aligned} \tag{2.8}$$

The face of the tetrahedron with the normal $-\boldsymbol{e}_i$ will be denoted by $\partial\omega^{(\boldsymbol{e}_i)}$, $i = 1, 2, 3$. Let $\boldsymbol{t}^{(\boldsymbol{n})}$ be the Cauchy stress vector at $\partial\omega^{(\boldsymbol{n})}$. Assuming that the Cauchy stress vectors depend only on the normal of the respective face, they are constant on

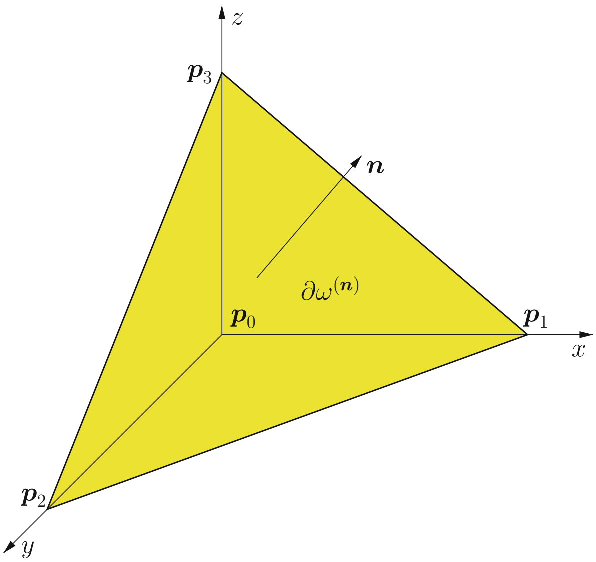

Fig. 2.1 Illustration of the tetrahedron used for discussing the linear dependency of the Cauchy stress vector on the normal

each face of the tetrahedron and the integrals on the faces can be computed easily. Applying in addition Newton's second law of motion (2.5) and the formula for the volume of a tetrahedron leads to

$$\underbrace{\boldsymbol{t}^{(\boldsymbol{n})}\left|\partial\omega^{(\boldsymbol{n})}\right| - \sum_{i=1}^{3}\boldsymbol{t}^{(\boldsymbol{e}_i)}\left|\partial\omega^{(\boldsymbol{e}_i)}\right|}_{\text{internal force}} = \underbrace{\rho\frac{h^{(\boldsymbol{n})}}{3}\left|\partial\omega^{(\boldsymbol{n})}\right|}_{\text{mass}}\boldsymbol{a}, \quad [\mathrm{N}] \tag{2.9}$$

where $|\cdot|$ is the area of the faces, $\boldsymbol{t}^{(\boldsymbol{e}_i)}$ the constant stress vector at face $\partial\omega^{(\boldsymbol{e}_i)}$, $\boldsymbol{a}$ $[\mathrm{m/s^2}]$ is an acceleration, and $h^{(\boldsymbol{n})}$ is the distance of the face $\partial\omega^{(\boldsymbol{n})}$ to the origin. The area of $\partial\omega^{(\boldsymbol{n})}$ can be calculated with the cross product, giving

$$\left|\partial\omega^{(\boldsymbol{n})}\right| = \frac{1}{2}\left|(\boldsymbol{p}_2-\boldsymbol{p}_1)\times(\boldsymbol{p}_3-\boldsymbol{p}_1)\right| = \frac{1}{2}\left\|\begin{pmatrix} y_2z_3 \\ z_3x_1 \\ x_1y_2 \end{pmatrix}\right\|_2 .$$

Using the representation (2.8) of the normal leads to

$$\left|\partial\omega^{(\boldsymbol{e}_1)}\right| = \frac{1}{2}y_2z_3 = \frac{1}{2}n_1\left\|(\boldsymbol{p}_2-\boldsymbol{p}_1)\times(\boldsymbol{p}_3-\boldsymbol{p}_1)\right\|_2 = \left|\partial\omega^{(\boldsymbol{n})}\right| n_1.$$

Analogous formulas are derived for $\left|\partial\omega^{(\boldsymbol{e}_2)}\right|$ and $\left|\partial\omega^{(\boldsymbol{e}_3)}\right|$. Inserting these formulas into (2.9) gives

$$\boldsymbol{t}^{(\boldsymbol{n})} - \sum_{i=1}^{3}\boldsymbol{t}^{(\boldsymbol{e}_i)}n_i = \rho\frac{h^{(\boldsymbol{n})}}{3}\boldsymbol{a}. \tag{2.10}$$

Shrinking now the tetrahedron to the origin, where $\partial\omega^{(\boldsymbol{n})}$ moves in the direction $\boldsymbol{n}$, the left-hand side of (2.10) stays constant whereas the right-hand side vanishes since $h^{(\boldsymbol{n})} \to 0$. Hence, one obtains in the limit that

$$\boldsymbol{t}^{(\boldsymbol{n})} = \sum_{i=1}^{3}\boldsymbol{t}^{(\boldsymbol{e}_i)}n_i = \left(\boldsymbol{t}^{(\boldsymbol{e}_1)}\ \boldsymbol{t}^{(\boldsymbol{e}_2)}\ \boldsymbol{t}^{(\boldsymbol{e}_3)}\right)\boldsymbol{n},$$

where $(\cdot,\cdot,\cdot)$ denotes a tensor with the respective columns. This relation means that $\boldsymbol{t}^{(\boldsymbol{n})}$ depends linearly on $\boldsymbol{n}$.

Thus, the model for the Cauchy stress vector is

$$\boldsymbol{t} = \mathbb{S}\boldsymbol{n}, \tag{2.11}$$

where $\mathbb{S}(t,\boldsymbol{x})$ $[\mathrm{N/m^2}]$ is a 3×3-tensor that is called stress tensor. The stress tensor represents all internal forces of the flow. Inserting (2.11) in the term representing the internal forces in (2.7) and applying the divergence theorem gives

$$\int_{\partial\omega} \boldsymbol{t}(t,\boldsymbol{s})\, d\boldsymbol{s} = \int_{\omega} \nabla\cdot\mathbb{S}(t,\boldsymbol{x})\, d\boldsymbol{x},$$

where the divergence of a tensor is defined row-wise

$$\nabla\cdot\mathbb{A} = \begin{pmatrix} \partial_x a_{11} + \partial_y a_{12} + \partial_z a_{13} \\ \partial_x a_{21} + \partial_y a_{22} + \partial_z a_{23} \\ \partial_x a_{31} + \partial_y a_{32} + \partial_z a_{33} \end{pmatrix}.$$

Since (2.7) holds for every volume ω, it follows that

$$\rho\left(\partial_t \boldsymbol{v} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v}\right) = \nabla\cdot\mathbb{S} + \boldsymbol{f}_{\text{ext}} \quad \forall\, t\in(0,T],\ \boldsymbol{x}\in\Omega. \tag{2.12}$$

This relation is the momentum equation. □

Remark 2.9 (Symmetry of the Stress Tensor) Let ω be an arbitrary volume with sufficiently smooth boundary $\partial\omega$ and let the net force be given by the right-hand side of (2.7). The torque in ω with respect to the origin $\mathbf{0}$ of the coordinate system is then defined by

$$\boldsymbol{M}_{\mathbf{0}} = \int_{\omega} \boldsymbol{r}\times\boldsymbol{f}_{\text{ext}}\, d\boldsymbol{x} + \int_{\partial\omega} \boldsymbol{r}\times(\mathbb{S}\boldsymbol{n})\ d\boldsymbol{s} \quad [\text{Nm}], \tag{2.13}$$

where (2.11) was used. In (2.13), $\boldsymbol{r} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3$ is the vector pointing from $\mathbf{0}$ to a point $\boldsymbol{x}\in\overline{\omega}$. A straightforward calculation shows that

$$\boldsymbol{r}\times(\mathbb{S}\boldsymbol{n}) = (\boldsymbol{r}\times\mathbb{S}_{*1}\ \boldsymbol{r}\times\mathbb{S}_{*2}\ \boldsymbol{r}\times\mathbb{S}_{*3})\,\boldsymbol{n},$$

where $\mathbb{S}_{*i}$ is the i-th column of $\mathbb{S}$ and $(\cdot)$ denotes here the tensor with the respective columns. Inserting this expression in (2.13), applying integration by parts, and using the product rule leads to

$$\begin{aligned} \boldsymbol{M}_{\mathbf{0}} &= \int_{\omega} \boldsymbol{r}\times\boldsymbol{f}_{\text{ext}}\, d\boldsymbol{x} + \int_{\omega} \nabla\cdot\left((\boldsymbol{r}\times\mathbb{S}_{*1}\ \boldsymbol{r}\times\mathbb{S}_{*2}\ \boldsymbol{r}\times\mathbb{S}_{*3})\right) d\boldsymbol{x} \\ &= \int_{\omega} \boldsymbol{r}\times(\boldsymbol{f}_{\text{ext}} + \nabla\cdot\mathbb{S})\, d\boldsymbol{x} \\ &\quad + \int_{\omega} \partial_x\boldsymbol{r}\times\mathbb{S}_{*1} + \partial_y\boldsymbol{r}\times\mathbb{S}_{*2} + \partial_z\boldsymbol{r}\times\mathbb{S}_{*3}\, d\boldsymbol{x}. \end{aligned} \tag{2.14}$$

Consider now a fluid in equilibrium state, i.e., the net forces acting on this fluid are zero. Hence, the right-hand side of (2.12) vanishes and so the first integral of (2.14). In addition, equilibrium requires in particular that $\boldsymbol{M}_{\mathbf{0}} = \mathbf{0}$. Thus, from (2.14) it follows that

$$\mathbf{0} = \int_{\omega} \partial_x \boldsymbol{r} \times \mathbb{S}_{*1} + \partial_y \boldsymbol{r} \times \mathbb{S}_{*2} + \partial_z \boldsymbol{r} \times \mathbb{S}_{*3} \, d\boldsymbol{x}. \tag{2.15}$$

Using now

$$\partial_x \boldsymbol{r} = \lim_{\Delta x \to 0} \frac{(x + \Delta x)\boldsymbol{e}_1 - x\boldsymbol{e}_1}{\Delta x} = \boldsymbol{e}_1,$$

$\partial_y \boldsymbol{r} = \boldsymbol{e}_2$, $\partial_z \boldsymbol{r} = \boldsymbol{e}_3$, and inserting these equations in (2.15) leads finally to

$$\mathbf{0} = \int_{\omega} \begin{pmatrix} \mathbb{S}_{32} - \mathbb{S}_{23} \\ \mathbb{S}_{13} - \mathbb{S}_{31} \\ \mathbb{S}_{21} - \mathbb{S}_{12} \end{pmatrix} (t, \boldsymbol{x}) \, d\boldsymbol{x}$$

for an arbitrary volume ω. From this relation, one deduces that $\mathbb{S}$ has to be symmetric, $\mathbb{S} = \mathbb{S}^T$, and $\mathbb{S}$ possesses six unknown components. □

Remark 2.10 (Decomposition of the Stress Tensor) To model the stress tensor in the basic variables introduced in Remark 2.1, this tensor is decomposed into

$$\mathbb{S} = \mathbb{V} - P\mathbb{I}. \tag{2.16}$$

Here, $\mathbb{V}$ [N/m²] is the so-called viscous stress tensor, representing the forces coming from the friction of the particles, and P [Pa] is the pressure, describing the forces acting on the surface of each fluid volume ω, where $\mathbb{I}$ is the identity tensor. The viscous stress tensor will be modeled in terms of the velocity, see Remark 2.12. □

Remark 2.11 (The Pressure) The pressure P acts on a surface of a fluid volume ω only normal to that surface and it is directed into ω. This property is reflected by the negative sign in the ansatz (2.16) since

$$-\int_{\partial\omega} P\boldsymbol{n} \, ds = -\int_{\omega} \nabla P \, d\boldsymbol{x} = -\int_{\omega} \nabla \cdot (P\mathbb{I}) \, d\boldsymbol{x}.$$

□

Remark 2.12 (The Viscous Stress Tensor) Friction between fluid particles can only occur if the particles move with different velocities. For this reason, the viscous stress tensor is modeled to depend on the gradient of the velocity. For the reason of symmetry, Remark 2.9, it is modeled to depend on the symmetric part of the gradient, the so-called velocity rate-of-deformation tensor or shortly velocity deformation tensor

$$\mathbb{D}(\boldsymbol{v}) = \frac{\nabla \boldsymbol{v} + (\nabla \boldsymbol{v})^T}{2} \quad [1/\mathrm{s}].$$

The gradient of the velocity is a tensor with the components

$$(\nabla \boldsymbol{v})_{ij} = \partial_j v_i = \frac{\partial v_i}{\partial x_j}, \quad i, j = 1, 2, 3.$$

If the velocity gradients are not too large, one can assume that first the dependency is linear and second that higher order derivatives can be neglected. Since there is no friction for a flow with constant velocity, such that $\mathbb{V}$ vanishes in this case, lower order terms than first order derivatives of the velocity should not appear in the model. The most general form of a tensor that satisfies all conditions is

$$\mathbb{V} = a\mathbb{D}(\boldsymbol{v}) + b(\nabla \cdot \boldsymbol{v})\mathbb{I},$$

where a and b do not depend on the velocity. Introducing the first order viscosity μ [$^{\mathrm{kg}}/_{(\mathrm{m\ s})}$] and the second order viscosity ζ [$^{\mathrm{kg}}/_{(\mathrm{m\ s})}$], one writes this tensor in fluid dynamics in the form

$$\mathbb{V} = 2\mu\mathbb{D}(\boldsymbol{v}) + \left(\zeta - \frac{2\mu}{3}\right)(\nabla \cdot \boldsymbol{v})\mathbb{I} \quad [^{\mathrm{N}}/_{\mathrm{m}^2}]. \tag{2.17}$$

The viscosity μ is also called dynamic or shear viscosity. The law (2.17) is for fluids the analog of Hooke's law for solids. □

Example 2.13 (Steady Rotation) There is no viscous stress, i.e., $\mathbb{V} = \mathbb{0}$, if the fluid is rotating steadily. In this situation, the velocity is given by

$$\boldsymbol{v} = \boldsymbol{\omega} \times \boldsymbol{x} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \omega_3 y - \omega_2 z \\ \omega_1 z - \omega_3 x \\ \omega_2 x - \omega_1 y \end{pmatrix},$$

where $\boldsymbol{\omega}$ [$^1/_{\mathrm{s}}$] is a constant angular velocity. One has obviously $\nabla \cdot \boldsymbol{v} = 0$ and

$$\nabla \boldsymbol{v} = \begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix} \quad \Longrightarrow \quad \mathbb{D}(\boldsymbol{v}) = \mathbb{0}.$$

Hence, (2.17) is an appropriate model in this case. □

Remark 2.14 (Newtonian Fluids) The linear relation (2.17) is only an approximation for a real fluid. In general, the relation will be nonlinear. Only for small stresses, a linear approximation of the general stress-deformation relation can be used. A linear stress-deformation relation was postulate by Newton. For this reason, a fluid satisfying assumption (2.17) is called Newtonian fluid. More general relations than (2.17) exist, however they are less well understood from the mathematical point of view. □

Remark 2.15 (Normal and Shear Stresses, Trace of the Stress Tensor) The diagonal components $\mathbb{S}_{11}, \mathbb{S}_{22}, \mathbb{S}_{33}$ of the stress tensor are called normal stresses and the off-diagonal components shear stresses.

For incompressible flows one gets with (2.4), (2.16), and (2.17)

$$\mathbb{S} = 2\mu\mathbb{D}(\boldsymbol{v}) - P\mathbb{I}. \tag{2.18}$$

The trace of the stress tensor is the sum of the normal stresses

$$\begin{aligned}\operatorname{tr}(\mathbb{S}) &= \mathbb{S}_{11} + \mathbb{S}_{22} + \mathbb{S}_{33} \\ &= 2\mu(\partial_x v_1 + \partial_y v_2 + \partial_z v_3) + 3\left(\zeta - \frac{2\mu}{3}\right)(\nabla\cdot\boldsymbol{v}) - 3P \\ &= 3\zeta(\nabla\cdot\boldsymbol{v}) - 3P.\end{aligned}$$

For incompressible fluids, it follows that

$$\operatorname{tr}(\mathbb{D}(\boldsymbol{v})) = \frac{1}{2\mu}(\operatorname{tr}(\mathbb{S}) + \operatorname{tr}(P\mathbb{I})) = \frac{1}{2\mu}(-3P + 3P) = 0$$

and

$$P(t,\boldsymbol{x}) = -\frac{1}{3}(\mathbb{S}_{11} + \mathbb{S}_{22} + \mathbb{S}_{33})(t,\boldsymbol{x}). \tag{2.19}$$

□

Remark 2.16 (The Navier–Stokes Equations) Now, the pressure part of the stress tensor and the model (2.17) of the viscous stress tensor can be inserted in (2.12) giving the general Navier–Stokes equations (including the conservation of mass)

$$\begin{aligned}\rho(\partial_t\boldsymbol{v} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v}) - 2\nabla\cdot(\mu\mathbb{D}(\boldsymbol{v})) \\ -\nabla\cdot\left(\left(\zeta - \frac{2\mu}{3}\right)(\nabla\cdot\boldsymbol{v})\mathbb{I}\right) + \nabla P = \boldsymbol{f}_{\text{ext}} \text{ in } (0,T]\times\Omega, \\ \partial_t\rho + \nabla\cdot(\rho\boldsymbol{v}) = 0 \quad \text{in } (0,T]\times\Omega.\end{aligned} \tag{2.20}$$

If the fluid is incompressible and homogeneous, such that μ and ρ are positive constants, the Navier–Stokes equations simplify to

$$\begin{aligned}\partial_t\boldsymbol{v} - 2\nu\nabla\cdot\mathbb{D}(\boldsymbol{v}) + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v} + \nabla\frac{P}{\rho} = \frac{\boldsymbol{f}_{\text{ext}}}{\rho} \text{ in } (0,T]\times\Omega, \\ \nabla\cdot\boldsymbol{v} = 0 \quad \text{in } (0,T]\times\Omega.\end{aligned} \tag{2.21}$$

Here, $\nu = \mu/\rho$ [m^2/s] is the kinematic viscosity of the fluid. □

2.3 The Dimensionless Navier–Stokes Equations

Remark 2.17 (Characteristic Scales) Mathematical analysis and numerical simulations are based on dimensionless equations. To derive dimensionless equations from system (2.21), the quantities

- L [m]—a characteristic length scale of the flow problem,
- U [m/s]—a characteristic velocity scale of the flow problem,
- T^* [s]—a characteristic time scale of the flow problem,

are introduced. □

Remark 2.18 (The Navier–Stokes Equations in Dimensionless Form) Denote by $(t', \boldsymbol{x}')$ [s, m] the old variables. Applying the transform of variables

$$\boldsymbol{x} = \frac{\boldsymbol{x}'}{L}, \quad \boldsymbol{u} = \frac{\boldsymbol{v}}{U}, \quad t = \frac{t'}{T^*}, \tag{2.22}$$

one obtains from (2.21) and a rescaling

$$\begin{aligned} \frac{L}{UT^*}\partial_t \boldsymbol{u} - \frac{2\nu}{UL}\nabla\cdot\mathbb{D}(\boldsymbol{u}) + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla\frac{P}{\rho U^2} &= \frac{L}{\rho U^2}\boldsymbol{f}_{\text{ext}} && \text{in } (0,T]\times\Omega,\\ \nabla\cdot\boldsymbol{u} &= 0 && \text{in } (0,T]\times\Omega, \end{aligned}$$

where all derivatives are with respect to the new variables. Without having emphasized this issue in the notation, also the domain and the time interval are now dimensionless. Defining

$$p = \frac{P}{\rho U^2}, \quad \text{Re} = \frac{UL}{\nu}, \quad \text{St} = \frac{L}{UT^*}, \quad \boldsymbol{f} = \frac{L}{\rho U^2}\boldsymbol{f}_{\text{ext}}, \tag{2.23}$$

the incompressible Navier–Stokes equations in dimensionless form

$$\begin{aligned} \text{St}\partial_t \boldsymbol{u} - \frac{2}{\text{Re}}\nabla\cdot\mathbb{D}(\boldsymbol{u}) + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p &= \boldsymbol{f} \quad \text{in } (0,T]\times\Omega,\\ \nabla\cdot\boldsymbol{u} &= 0 \quad \text{in } (0,T]\times\Omega, \end{aligned} \tag{2.24}$$

are obtained. The constant Re is called Reynolds number and the constant St Strouhal number. These numbers allow the classification and comparison of different flows. □

Remark 2.19 (Inherent Difficulties of the Dimensionless Navier–Stokes Equations) To simplify the notations, one uses the characteristic time scale $T^* = L/U$ such that (2.24) simplifies to

$$\begin{aligned} \partial_t \boldsymbol{u} - 2\nu\nabla\cdot\mathbb{D}(\boldsymbol{u}) + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p &= \boldsymbol{f} \text{ in } (0,T]\times\Omega,\\ \nabla\cdot\boldsymbol{u} &= 0 \text{ in } (0,T]\times\Omega, \end{aligned} \tag{2.25}$$

with the dimensionless viscosity $\nu = \mathrm{Re}^{-1}$. Here, with an abuse of notation, the same symbol is used as for the kinematic viscosity.

This transform and the resulting Eq. (2.25) are the basic equations for the mathematical analysis of the incompressible Navier–Stokes equations and the numerical simulation of incompressible flows. System (2.25) comprises two important difficulties:

- the coupling of velocity and pressure,
- the nonlinearity of the convective term.

Additionally, difficulties for the numerical simulation occur if

- the convective term dominates the viscous term, i.e., if ν is small.

□

Remark 2.20 (Different Forms of Terms in (2.25)*)* With the help of the divergence constraint, i.e., the second equation in (2.25), the viscous and the convective term of the Navier–Stokes equations can be reformulated equivalently.

Assume that $\boldsymbol{u}$ is sufficiently smooth with $\nabla \cdot \boldsymbol{u} = 0$. Then, straightforward calculations, using the Theorem of Schwarz and the second equation of (2.25), give

$$\nabla \cdot (\nabla \boldsymbol{u}) = \Delta \boldsymbol{u}, \quad \nabla \cdot \left(\nabla \boldsymbol{u}^T\right) = \nabla (\nabla \cdot \boldsymbol{u}) = \begin{pmatrix} \partial_x (\nabla \cdot \boldsymbol{u}) \\ \partial_y (\nabla \cdot \boldsymbol{u}) \\ \partial_z (\nabla \cdot \boldsymbol{u}) \end{pmatrix} = \boldsymbol{0}. \tag{2.26}$$

Thus, the viscous term becomes

$$-2\nu \nabla \cdot \mathbb{D}(\boldsymbol{u}) = -\nu \Delta \boldsymbol{u}. \tag{2.27}$$

For the convective term, ones uses the identity (product rule)

$$\begin{aligned} \nabla \cdot \left(\boldsymbol{u}\boldsymbol{v}^T\right) &= \begin{pmatrix} \partial_x (u_1 v_1) + \partial_y (u_1 v_2) + \partial_z (u_1 v_3) \\ \partial_x (u_2 v_1) + \partial_y (u_2 v_2) + \partial_z (u_2 v_3) \\ \partial_x (u_3 v_1) + \partial_y (u_3 v_2) + \partial_z (u_3 v_3) \end{pmatrix} \\ &= \begin{pmatrix} u_1 \left(\partial_x v_1 + \partial_y v_2 + \partial_z v_3\right) \\ u_2 \left(\partial_x v_1 + \partial_y v_2 + \partial_z v_3\right) \\ u_3 \left(\partial_x v_1 + \partial_y v_2 + \partial_z v_3\right) \end{pmatrix} + \begin{pmatrix} v_1 \partial_x u_1 + v_2 \partial_y u_1 + v_3 \partial_z u_1 \\ v_1 \partial_x u_2 + v_2 \partial_y u_2 + v_3 \partial_z u_2 \\ v_1 \partial_x u_3 + v_2 \partial_y u_3 + v_3 \partial_z u_3 \end{pmatrix} \\ &= (\nabla \cdot \boldsymbol{v})\boldsymbol{u} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{u}. \end{aligned} \tag{2.28}$$

In the case $\boldsymbol{v} = \boldsymbol{u}$ with $\nabla \cdot \boldsymbol{u} = 0$, it follows that

$$(\boldsymbol{u} \cdot \nabla)\, \boldsymbol{u} = \nabla \cdot \left(\boldsymbol{u}\boldsymbol{u}^T\right). \tag{2.29}$$

A detailed presentation and discussion of different forms of the convective term is given in Sect. 6.1.2. □

Remark 2.21 (Two-dimensional Navier–Stokes Equations) Even if real flows occur only in three dimensions, the consideration of the Navier–Stokes equations (2.25) in two dimensions is also of interest. There are applications where the flow is constant in the third direction and it behaves virtually two-dimensional. □

Remark 2.22 (Special Cases of Incompressible Flow Models)

- In a stationary flow, the velocity and the pressure do not change in time. Hence $\partial_t \boldsymbol{u} = \boldsymbol{0}$ and these flows are modeled by the so-called stationary or steady-state Navier–Stokes equations

$$\begin{aligned} -\nu \Delta \boldsymbol{u} + (\boldsymbol{u} \cdot \nabla)\boldsymbol{u} + \nabla p &= \boldsymbol{f} \text{ in } \Omega, \\ \nabla \cdot \boldsymbol{u} &= 0 \text{ in } \Omega. \end{aligned} \tag{2.30}$$

  A necessary condition for the time-independence of a flow field is that the data of the problem, i.e., the right-hand side and the boundary conditions, see Sect. 2.4, are time-independent. But this condition is not sufficient, cf. Example D.8.
- If in a stationary flow the viscous transport dominates the convective transport, i.e., if the fluid moves very slowly, the nonlinear convective term of the Navier–Stokes equations (2.30) can be neglected. This situation leads to a linear system of equations, the so-called Stokes equations

$$\begin{aligned} -\Delta \boldsymbol{u} + \nabla p &= \boldsymbol{f} \text{ in } \Omega, \\ \nabla \cdot \boldsymbol{u} &= 0 \text{ in } \Omega. \end{aligned} \tag{2.31}$$

  Here, the momentum equation was divided by ν, defining a new pressure and a new right-hand side.
- In some standard schemes for solving the Navier–Stokes equations numerically, the so-called Oseen equations appear. Given a divergence-free flow field $\boldsymbol{b}$, the Oseen equations are a system of linear equations of the form

$$\begin{aligned} -\nu \Delta \boldsymbol{u} + (\boldsymbol{b} \cdot \nabla)\boldsymbol{u} + \nabla p + c\boldsymbol{u} &= \boldsymbol{f} \text{ in } \Omega, \\ \nabla \cdot \boldsymbol{u} &= 0 \text{ in } \Omega, \end{aligned} \tag{2.32}$$

  with a scalar-valued function $c(\boldsymbol{x}) \geq 0$.

□

2.4 Initial and Boundary Conditions

Remark 2.23 (General Considerations) The Navier–Stokes equations (2.25) are a first order partial differential equation with respect to time and a second order partial differential equation with respect to space. Thus, they have to be equipped with an initial condition for the velocity at $t = 0$ and with boundary conditions on the boundary $\Gamma = \partial\Omega$ of Ω, if Ω is a bounded domain. There are several kinds of boundary conditions which can be prescribed for incompressible flows. Of course, a compatibility condition should be fulfilled between the boundary conditions of the initial velocity field and the limit of the prescribed boundary conditions for $t \to 0$, $t > 0$.

In applications, the initial and boundary conditions are given in terms of quantities with dimensions. For the analysis and the simulation of the Navier–Stokes equations, these conditions have to be converted to non-dimensional quantities with the characteristic scales from Remark 2.17. Here, only conditions for the dimensionless equations will be discussed. □

Remark 2.24 (Initial Condition) Concerning the initial condition, an initial velocity field $\boldsymbol{u}(0, \boldsymbol{x}) = \boldsymbol{u}_0(\boldsymbol{x})$ is prescribed at $t = 0$. The initial flow field has to be in some sense divergence-free. □

Remark 2.25 (Dirichlet Boundary Conditions, No-slip Boundary Conditions, Essential Boundary Conditions) Often used boundary conditions describe the velocity field at a part of the boundary

$$\boldsymbol{u}(t, \boldsymbol{x}) = \boldsymbol{g}(t, \boldsymbol{x}) \quad \text{in} \quad (0, T] \times \Gamma_{\text{diri}},$$

with $\Gamma_{\text{diri}} \subset \Gamma$. This boundary condition is called Dirichlet boundary condition. It models in particular prescribed inflows into Ω and outflows from Ω.

In the special case $\boldsymbol{g}(t, \boldsymbol{x}) = \boldsymbol{0}$ in $(0, T] \times \Gamma_{\text{diri}}$, this boundary condition is called no-slip boundary condition. The no-slip condition is usually applied at fixed walls. Let $\boldsymbol{n}$ be the unit normal vector in $\boldsymbol{x} \in \Gamma_{\text{nosl}} \subset \Gamma_{\text{diri}}$ and $\{\boldsymbol{t}_1, \boldsymbol{t}_2\}$ unit tangential vectors such that $\{\boldsymbol{n}, \boldsymbol{t}_1, \boldsymbol{t}_2\}$ is an orthonormal system of vectors. Then, the no-slip boundary condition can be decomposed into three parts:

$$\boldsymbol{u}(t, \boldsymbol{x}) = \boldsymbol{0} \iff \boldsymbol{u}(t, \boldsymbol{x}) \cdot \boldsymbol{n} = 0, \ \boldsymbol{u}(t, \boldsymbol{x}) \cdot \boldsymbol{t}_1 = 0, \ \ \boldsymbol{u}(t, \boldsymbol{x}) \cdot \boldsymbol{t}_2 = 0$$

in $\boldsymbol{x} \in \Gamma_{\text{nosl}}$. The condition $\boldsymbol{u}(t, \boldsymbol{x}) \cdot \boldsymbol{n} = 0$ states that the fluid does not penetrate the wall. The other two conditions describe that the fluid does not slip along the wall.

If Dirichlet boundary conditions are prescribed on the whole boundary of Ω, there are two additional issues. First, the pressure is determined only up to an additive constant. An additional condition for fixing the constant has to be introduced, e.g., that the integral mean value of the pressure should vanish

$$\int_\Omega p(t, \boldsymbol{x}) \, d\boldsymbol{x} = 0 \quad t \in (0, T].$$

Second, from the divergence-free constraint and integration by parts it follows that the boundary condition has to satisfy the compatibility condition

$$0 = \int_\Omega \nabla \cdot \boldsymbol{u}(t, \boldsymbol{x}) \, d\boldsymbol{x} = \int_\Gamma (\boldsymbol{u} \cdot \boldsymbol{n}) \, (t, \boldsymbol{s}) \, d\boldsymbol{s} = \int_\Gamma (\boldsymbol{g} \cdot \boldsymbol{n})(t, \boldsymbol{s}) \, d\boldsymbol{s} \qquad (2.33)$$

for all times.

In the case of the Navier–Stokes equations and their special cases, Dirichlet boundary conditions are so-called essential boundary conditions. Such boundary conditions enter the definition of appropriate function spaces for the study of the equations in the framework of functional analysis, see Sect. 3.2. □

Remark 2.26 (Free Slip Boundary Conditions, Slip with Friction Boundary Conditions) The free slip boundary condition is applied on boundaries without friction. It has the form

$$\begin{aligned} \boldsymbol{u}\cdot\boldsymbol{n} &= g \quad \text{in } (0,T]\times\Gamma_{\text{slip}}, \\ \boldsymbol{n}^T\mathbb{S}\boldsymbol{t}_k &= 0 \quad \text{in } (0,T]\times\Gamma_{\text{slip}}, \quad 1\le k\le d-1, \end{aligned} \tag{2.34}$$

$\Gamma_{\text{slip}} \subset \Gamma$, where the dimensionless stress tensor is given by

$$\mathbb{S} = 2\nu\mathbb{D}(\boldsymbol{u}) - p\mathbb{I}. \tag{2.35}$$

There is no penetration through the wall if $g = 0$ on Γ_{slip}.

The slip with linear friction and no penetration boundary condition has the form

$$\begin{aligned} \boldsymbol{u}\cdot\boldsymbol{n} &= 0 \quad \text{in } (0,T]\times\Gamma_{\text{slfr}}, \\ \boldsymbol{u}\cdot\boldsymbol{t}_k + \beta^{-1}\boldsymbol{n}^T\mathbb{S}\boldsymbol{t}_k &= 0 \quad \text{in } (0,T]\times\Gamma_{\text{slfr}}, \quad 1\le k\le d-1, \end{aligned} \tag{2.36}$$

with $\Gamma_{\text{slfr}} \subset \Gamma$. This boundary condition states that the fluid does not penetrate the wall and it slips along the wall while loosing energy. The loss of energy is given by the friction parameter β. In the limit case $\beta^{-1} \to 0$, the no-slip condition is recovered and in the limit case $\beta^{-1} \to \infty$ the free slip condition. Slip with friction boundary conditions were studied already by Maxwell (1879) and Navier (1823). The difficulty in the application of this boundary condition consists in the determination of the friction parameter β, which might depend, e.g., on the roughness of the wall.

Since $\boldsymbol{n}$ and $\boldsymbol{t}_k$ are orthogonal vectors, the values of the pressure do not play any role in the boundary conditions (2.34) and (2.36). Hence, an additional condition for the pressure is needed to fix the additive constant. □

Remark 2.27 (Do-Nothing Boundary Condition, Natural Boundary Conditions) For numerical simulations, the so-called do-nothing boundary condition

$$\mathbb{S}\boldsymbol{n} = \boldsymbol{0} \;\text{ in }\; (0,T]\times\Gamma_{\text{donot}}, \tag{2.37}$$

$\Gamma_{\text{donot}} \subset \Gamma$ is often applied. This boundary condition models the situation that the normal stress, which is equal to the Cauchy stress vector (2.11), vanishes on the boundary part Γ_{donot}. A do-nothing boundary condition is often used if no other boundary condition at the outlet is available.

From the mathematical point of view, the do-nothing boundary condition is a natural boundary conditions. Deriving from the strong form of the equations (2.25) a so-called weak form, natural boundary conditions appear in the arising integrals on the boundary. In the special case of the do-nothing boundary condition, the integral on Γ_{donot} vanishes since the term in the integral is zero.

The boundary condition (2.37) contains also a contribution from the pressure. This issue fixes the problem of the additive constant, i.e., if on a part of the boundary the do-nothing boundary condition is prescribed, it is not necessary to introduce an additional condition for the pressure.

However, there are two problems with the do-nothing boundary condition (2.37):

- It turns out that this boundary condition is incorrect for a simple two-dimensional channel flow, see Example 2.28. This problem can be fixed by modifying the tensor in the boundary condition, compare (2.40) below.

  In practice, a slightly incorrect boundary condition at the outlet might be of minor importance. If possible, the computational domain can be extended such that the impact of the boundary condition at the outlet on the solution in regions of interest becomes negligible.
- For problems where the do-nothing boundary condition (2.37) or the modified do-nothing condition (2.40) is prescribed on a part of the boundary, the stability of the solution can be proved only with some additional assumption on the solution, see Remark 2.29.

□

Example 2.28 (Do-Nothing Conditions for a Two-dimensional Channel Flow Problem (Hagen–Poiseuille Flow)) Let $\Omega = (0, l_x) \times (-l_y, l_y)$, $l_x, l_y > 0$, be a bounded domain. On the boundary $x = 0$ the parabolic inflow condition

$$\boldsymbol{u} = U_{\text{in}} \begin{pmatrix} l_y^2 - y^2 \\ 0 \end{pmatrix}, \quad U_{\text{in}} > 0, \tag{2.38}$$

and at the boundaries $y = -l_y, y = l_y$, the no-slip condition $\boldsymbol{u} = \boldsymbol{0}$ are prescribed. There are no body forces in this problem, i.e., $\boldsymbol{f} = \boldsymbol{0}$, and the kinematic viscosity ν is assumed to be sufficiently large. Taking (2.38) as velocity of the flow field, one finds with direct calculations that

$$\partial_t \boldsymbol{u} = \boldsymbol{0}, \quad \nabla \cdot \boldsymbol{u} = 0, \quad (\boldsymbol{u} \cdot \nabla)\boldsymbol{u} = \boldsymbol{0}, \quad \Delta \boldsymbol{u} = 2\nabla \cdot \mathbb{D}(\boldsymbol{u}) = \begin{pmatrix} -2U_{\text{in}} \\ 0 \end{pmatrix}.$$

From (2.25) it follows that (2.38) together with

$$p = 2\nu U_{\text{in}}(x + C), \quad C \in \mathbb{R}, \tag{2.39}$$

is a solution of the Navier–Stokes equations. This solution is called Hagen–Poiseuille flow.

Now, the constant in the pressure (2.39) should be determined such that at the outlet $x = l_x$ the do-nothing condition (2.37) is satisfied. At the outlet it is $\boldsymbol{n} = (1, 0)^T$. One finds that

$$\mathbb{S}\boldsymbol{n} = (2\nu\mathbb{D}(\boldsymbol{u}) - p\mathbb{I})\,\boldsymbol{n} = 2\nu U_{\text{in}}\left[\begin{pmatrix} 0 & -y \\ -y & 0 \end{pmatrix} - \begin{pmatrix} x+C & 0 \\ 0 & x+C \end{pmatrix}\right]\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
$$= -2\nu U_{\text{in}}\begin{pmatrix} x+C \\ y \end{pmatrix}.$$

This expression does not vanish because the second component does not vanish. Hence, the do-nothing boundary condition (2.37) is not satisfied for the Hagen–Poiseuille flow.

A modification of the do-nothing boundary condition (2.37) consists in replacing the velocity deformation tensor with the velocity gradient, using (2.27),

$$(\nu\nabla\boldsymbol{u} - p\mathbb{I})\,\boldsymbol{n} = \boldsymbol{0} \ \text{ in } \ (0, T] \times \Gamma_{\text{donot}}, \tag{2.40}$$

$\Gamma_{\text{donot}} \subset \Gamma$. For the Hagen–Poiseuille flow, one obtains

$$(\nu\nabla\boldsymbol{u} - p\mathbb{I})\,\boldsymbol{n} = 2\nu U_{\text{in}}\left[\begin{pmatrix} 0 & -y \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} x+C & 0 \\ 0 & x+C \end{pmatrix}\right]\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$
$$= -2\nu U_{\text{in}}\begin{pmatrix} x+C \\ 0 \end{pmatrix}.$$

With the choice $C = -l_x$, the do-nothing condition (2.40) is satisfied at the boundary $x = l_x$.

The differences of the do-nothing conditions (2.37) and (2.40) for the Hagen–Poiseuille flow were noted, e.g., in Heywood et al. (1996, Fig. 10). With the do-nothing condition (2.37), the velocity field at the outlet becomes directed to the boundary of the channel. □

Remark 2.29 (Directional Do-Nothing Condition) Do-nothing conditions of the form (2.37) or (2.40) are usually applied at outlets of the domain. However, at outlets there might be also some inflow, e.g., if a vortex crosses the outlet.

From the mathematical point of view, the stability of flows with inflows in combination with do-nothing boundary conditions can be controlled only under a smallness assumption for the size of the inflow, see Braack and Mucha (2014) for details. To overcome this problem, a directional do-nothing condition can be used, reading

$$\left.\begin{aligned} (\nu\nabla\boldsymbol{u} - p\mathbb{I})\,\boldsymbol{n} &= \boldsymbol{0} \text{ if } \boldsymbol{u}\cdot\boldsymbol{n} \ge 0, \\ (\nu\nabla\boldsymbol{u} - p\mathbb{I})\,\boldsymbol{n} - \frac{1}{2}(\boldsymbol{u}\cdot\boldsymbol{n})\boldsymbol{u} &= \boldsymbol{0} \text{ if } \boldsymbol{u}\cdot\boldsymbol{n} < 0, \end{aligned}\right\} \ \text{in } \ (0, T] \times \Gamma_{\text{dirdonot}}, \tag{2.41}$$

$\Gamma_{\text{dirdonot}} \subset \Gamma$. If there is no inflow, i.e., $\boldsymbol{u} \cdot \boldsymbol{n} \geq 0$, the directional condition (2.41) reduces to the do-nothing condition (2.40). □

Remark 2.30 (A Boundary Condition on the Pressure) A boundary condition of the form

$$\boldsymbol{u} \times \boldsymbol{n} = \boldsymbol{g}_{\boldsymbol{u}} \times \boldsymbol{n} \quad \text{in } (0, T] \times \Gamma_{\text{pres}}, \tag{2.42}$$

$$p + \frac{1}{2}\boldsymbol{u} \cdot \boldsymbol{u} = g_p \qquad \text{in } (0, T] \times \Gamma_{\text{pres}}, \tag{2.43}$$

with $\Gamma_{\text{pres}} \subset \Gamma$ and given functions $\boldsymbol{g}_{\boldsymbol{u}}, g_p$ can be applied for the Navier–Stokes equations, e.g., see Bernardi et al. (2015). In two dimensions, (2.42) represents a condition on the tangential velocity

$$\boldsymbol{u} \times \boldsymbol{n} = \begin{pmatrix} 0 \\ 0 \\ \boldsymbol{u} \cdot \boldsymbol{t} \end{pmatrix}, \quad \text{with} \quad \boldsymbol{t} = \begin{pmatrix} n_2 \\ -n_1 \end{pmatrix}.$$

The quantity on the left-hand side of (2.43) is called Bernoulli pressure.

For the Stokes equations, the second term on the left-hand side of (2.43) has to be removed. □

Remark 2.31 (Conditions for an Infinite Domain, Periodic Boundary Conditions) The case $\Omega = \mathbb{R}^3$ is also considered in analytical and numerical studies of the Navier–Stokes equations. There are two situations in this case. In the first one, the decay of the velocity field as $\|\boldsymbol{x}\|_2 \to \infty$ is prescribed. The second situation consists of applying periodic boundary conditions. These boundary conditions do not possess any physical meaning. They are used to simulate an infinite extension of Ω in one or more directions, where it is assumed that the flow is periodic in this direction with the length l of the period. In computations, e.g., the cube $\Omega = (0, l)^d$ is used and the periodic boundary conditions are given by

$$\boldsymbol{u}(t, \boldsymbol{x} + l\boldsymbol{e}_i) = \boldsymbol{u}(t, \boldsymbol{x}) \quad \forall\ (t, \boldsymbol{x}) \in (0, T] \times \Gamma.$$

From the point of view of the finite computational domain, all appearing functions have to be extended periodically in the periodic direction to return to the original problem.

The use of space-periodic boundary conditions may also facilitate analytical investigations, see Temam (1995, p. 4). □

Chapter 3
Finite Element Spaces for Linear Saddle Point Problems

Remark 3.1 (Motivation) This chapter deals with the first difficulty inherent to the incompressible Navier–Stokes equations, see Remark 2.19, namely the coupling of velocity and pressure. The characteristic feature of this coupling is the absence of a pressure contribution in the continuity equation. In fact, the continuity equation can be considered as a constraint for the velocity and the pressure in the momentum equation as a Lagrangian multiplier. This kind of coupling is called saddle point problem.

Appropriate finite element spaces for velocity and pressure have to satisfy the so-called discrete inf-sup condition. This condition is derived on the basis of the theory for an abstract linear saddle point problem. Several techniques for proving the discrete inf-sup condition will be presented and applied for concrete pairs of finite element spaces for velocity and pressure.

All special cases of models for incompressible flow problems given in Remark 2.22 possess the same coupling of velocity and pressure, in particular the linear models of the Stokes and the Oseen equations. Linear problems are also of interest in the numerical simulation of the Navier–Stokes equations. After having discretized these equations implicitly in time, a nonlinear saddle point problem has to be solved in each discrete time. The solution of this problem is performed iteratively, requiring in each iteration step the solution of a linear saddle point problem for velocity and pressure. These linear saddle point problems will be discretized with finite element spaces. The existence and uniqueness of a solution of these discrete linear problems is crucial for performing the iteration. Altogether, the theory of linear saddle problems plays an essential role for the theory of all models for incompressible flows from Chap. 2.

A comprehensive presentation of the theory of linear saddle point problems can be found in the monograph Boffi et al. (2013). □

V. John, *Finite Element Methods for Incompressible Flow Problems*, Springer Series in Computational Mathematics 51, DOI 10.1007/978-3-319-45750-5_3

3.1 Existence and Uniqueness of a Solution of an Abstract Linear Saddle Point Problem

Remark 3.2 (Contents) This section presents an abstract framework for studying the existence and uniqueness of solutions of those types of linear saddle point problems which are of interest for incompressible flow problems. The presentation follows Girault and Raviart (1986, Chap. I, § 4). □

Remark 3.3 (Abstract Linear Saddle Point Problem) Let V and Q be two real Hilbert spaces with inner products $(\cdot,\cdot)_V$ and $(\cdot,\cdot)_Q$ and with induced norms $\|\cdot\|_V$ and $\|\cdot\|_Q$, respectively. Their corresponding dual spaces are given by V' and Q', with the dual pairing denoted by $\langle\cdot,\cdot\rangle_{V',V}$ and $\langle\cdot,\cdot\rangle_{Q',Q}$. The norms of the dual spaces are defined in the usual way by

$$\|\phi\|_{V'} := \sup_{v\in V, v\neq 0} \frac{\langle \phi, v\rangle_{V',V}}{\|v\|_V}, \quad \|\psi\|_{Q'} := \sup_{q\in Q, q\neq 0} \frac{\langle \psi, q\rangle_{Q',Q}}{\|q\|_Q}. \tag{3.1}$$

Two continuous bilinear forms are considered

$$a(\cdot,\cdot)\ :\ V\times V\to\mathbb{R}, \quad b(\cdot,\cdot)\ :\ V\times Q\to\mathbb{R}, \tag{3.2}$$

with the usual definition of their norms

$$\|a\| = \sup_{v,w\in V, v,w\neq 0} \frac{a(v,w)}{\|v\|_V\,\|w\|_V}, \quad \|b\| = \sup_{v\in V, q\in Q, v,q\neq 0} \frac{b(v,q)}{\|v\|_V\,\|q\|_Q}. \tag{3.3}$$

The following problem is studied: Find $(u,p)\in V\times Q$ such that for given $(f,r)\in V'\times Q'$

$$\begin{aligned} a(u,v) + b(v,p) &= \langle f, v\rangle_{V',V} \quad \forall\ v\in V, \\ b(u,q) &= \langle r, q\rangle_{Q',Q} \quad \forall\ q\in Q. \end{aligned} \tag{3.4}$$

System (3.4) is called linear saddle point problem. Concrete choices of the spaces and bilinear forms for incompressible flow problems are discussed in Sect. 3.2. □

Remark 3.4 (Operator Form of the Linear Saddle Point Problem) Problem (3.4) can be transformed into an equivalent form using operators instead of bilinear forms. Linear operators can be defined which are associated with the bilinear forms given in (3.2):

$$A\in\mathcal{L}\left(V,V'\right) \text{ defined by } \langle Au, v\rangle_{V',V} = a(u,v) \quad \forall\ u,v\in V,$$

$$B\in\mathcal{L}\left(V,Q'\right) \text{ defined by } \langle Bu, q\rangle_{Q',Q} = b(u,q) \quad \forall\ u\in V,\ \forall\ q\in Q.$$

Using the definition of the norms of the dual spaces (3.1), the norms of the operators are given by

$$\|Av\|_{V'} = \sup_{w\in V, w\neq 0} \frac{\langle Av, w\rangle_{V',V}}{\|w\|_V} \quad \Longrightarrow$$

$$\|A\|_{\mathcal{L}(V,V')} = \sup_{v\in V, v\neq 0} \frac{\|Av\|_{V'}}{\|v\|_V} = \sup_{v,w\in V, v,w\neq 0} \frac{a(v,w)}{\|v\|_V \|w\|_V} = \|a\| ,$$

and analogously

$$\|B\|_{\mathcal{L}(V,Q')} = \|b\| .$$

Let $B' \in \mathcal{L}(Q, V')$ be the adjoint (dual) operator of B defined by

$$\langle B'q, v\rangle_{V',V} = \langle Bv, q\rangle_{Q',Q} = b(v,q) \quad \forall\ v \in V,\ \forall\ q \in Q.$$

With these operators, Problem (3.4) can be written in the equivalent form: Find $(u, p) \in V \times Q$ such that

$$\begin{aligned} Au + B'p &= f \ \text{ in } V', \\ Bu &= r \ \text{ in } Q'. \end{aligned} \tag{3.5}$$

□

Definition 3.5 (Well-posedness of Problem (3.5)) Let

$$\Phi \in \mathcal{L}\left(V \times Q, V' \times Q'\right) \ : \ \Phi(v,q) = (Av + B'q, Bv)$$

be a linear operator, where $(\cdot,\cdot)$ denotes a vector with two components. Problem (3.5) is said to be well-posed if $\Phi(\cdot,\cdot)$ is an isomorphism from $V \times Q$ onto $V' \times Q'$. □

Remark 3.6 (On Definition 3.5) Definition 3.5 means that Problem (3.5) possesses for all possible right-hand sides a unique solution. The purpose of the following studies consists in deriving necessary and sufficient conditions for (3.5) to be well-posed. □

Remark 3.7 (The Finite-dimensional Case) Consider for the moment that V and Q are finite-dimensional spaces of dimension n_V and n_Q, respectively. Then, the operators in (3.5) can be represented with matrices, with $B' = B^T$, and the functions with vectors. The well-posedness of (3.5) means that the linear system of equations

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{r} \end{pmatrix}, \quad \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \in \mathbb{R}^{(n_V+n_Q)\times(n_V+n_Q)}, \tag{3.6}$$

has a unique solution or, equivalently, that the system matrix is non-singular. Here, conditions will be derived such that this property is given. These considerations should provide an idea of the kind of conditions to be expected in the general case.

Separate Consideration of Velocity and Pressure A possible way to solve (3.6) starts by solving the first equation of (3.6) for $\underline{u}$

$$\underline{u} = A^{-1}\left(\underline{f} - B^T\underline{p}\right). \tag{3.7}$$

Inserting this expression into the second equation gives

$$\left(BA^{-1}B^T\right)\underline{p} = BA^{-1}\underline{f} - \underline{r}. \tag{3.8}$$

If (3.8) possesses a unique solution $\underline{p}$, this solution can be inserted into (3.7) and a unique solution $\underline{u}$ is obtained, too. This way to compute a unique solution works if

- $A : V \to V'$ is an isomorphism, i.e., A is non-singular,
- $BA^{-1}B^T : Q \to Q'$ is an isomorphism, i.e., $BA^{-1}B^T$ is non-singular.

Let $\underline{p}$ be a solution of (3.8). Then, also $\underline{p} + \underline{\tilde{p}}$ with $\underline{\tilde{p}} \in \ker\left(B^T\right)$ is a solution of (3.8). Thus, for $BA^{-1}B^T$ to be non-singular, it is necessary that $\ker\left(B^T\right) = \{\underline{0}\}$ or equivalently that $B^T : Q \to V' = V$ is injective. With a similar argument, one finds that B must be injective on the range of $A^{-1}B^T$, i.e., $\ker(B) \cap \operatorname{range}\left(A^{-1}B^T\right) = \{\underline{0}\}$.

Joint Consideration of Velocity and Pressure One can also consider the system matrix (3.6) as a whole. A first necessary condition for the matrix to be non-singular is $n_Q \le n_V$, since the last rows of the system matrix span a space of dimension at most n_V (only the first n_V entries of these rows might be non-zero). Assume that A is non-singular, then the system matrix is non-singular if and only if B has full rank, i.e., $\operatorname{rank}(B) = n_Q$. It will be shown now that $\operatorname{rank}(B) = n_Q$ if and only if

$$\inf_{\underline{q}\in\mathbb{R}^{n_Q},\underline{q}\ne\underline{0}}\ \sup_{\underline{v}\in\mathbb{R}^{n_V},\underline{v}\ne\underline{0}} \frac{\underline{v}^T B^T \underline{q}}{\|\underline{v}\|_2 \left\|\underline{q}\right\|_2} \ge \beta > 0. \tag{3.9}$$

Let (3.9) be satisfied and let $\operatorname{rank}(B) < n_Q$. Then, there is a $\underline{q} \in \mathbb{R}^{n_Q}$, $\underline{q} \ne \underline{0}$, such that $\underline{q} \in \ker\left(B^T\right)$, i.e., $B^T\underline{q} = \underline{0}$. For this vector, it is $\underline{v}^T B^T \underline{q} = \underline{0}$ for all $\underline{v} \in \mathbb{R}^{n_V}$ such that the supremum of (3.9) is zero and (3.9) cannot be satisfied. This result is a contradiction and hence $\operatorname{rank}(B) = n_Q$.

On the other hand, let $\operatorname{rank}(B) = n_Q$. Then, for each $\underline{q} \in \mathbb{R}^{n_Q}$, $\underline{q} \neq \underline{0}$, one has that $B^T\underline{q} \neq \underline{0}$ with $B^T\underline{q} \in \mathbb{R}^{n_V}$. Choosing $\underline{v} = B^T\underline{q}$ gives

$$\begin{aligned}\inf_{\underline{q}\in\mathbb{R}^{n_Q},\underline{q}\neq\underline{0}} \sup_{\underline{v}\in\mathbb{R}^{n_V},\underline{v}\neq\underline{0}} \frac{\underline{v}^T B^T \underline{q}}{\|\underline{v}\|_2 \left\|\underline{q}\right\|_2} &\geq \inf_{\underline{q}\in\mathbb{R}^{n_Q},\underline{q}\neq\underline{0}} \frac{\left\|B^T\underline{q}\right\|_2^2}{\left\|B^T\underline{q}\right\|_2 \left\|\underline{q}\right\|_2} \\ &= \inf_{\underline{q}\in\mathbb{R}^{n_Q},\underline{q}\neq\underline{0}} \frac{\left\|B^T\underline{q}\right\|_2}{\left\|\underline{q}\right\|_2}. \end{aligned} \tag{3.10}$$

It is

$$\frac{\left\|B^T\underline{q}\right\|_2^2}{\left\|\underline{q}\right\|_2^2} = \frac{\underline{q}^T BB^T \underline{q}}{\underline{q}^T\underline{q}}.$$

This expression is a Rayleigh quotient and it is known that

$$\inf_{\underline{q}\in\mathbb{R}^{n_Q},\underline{q}\neq\underline{0}} \frac{\underline{q}^T BB^T \underline{q}}{\underline{q}^T\underline{q}} = \lambda_{\min}\left(BB^T\right),$$

where $\lambda_{\min}\left(BB^T\right)$ is the smallest eigenvalue of BB^T, see Lemma A.19. Since B was assumed to have full rank, one has $\lambda_{\min}\left(BB^T\right) > 0$ and hence with (3.10)

$$\inf_{\underline{q}\in\mathbb{R}^{n_Q},\underline{q}\neq\underline{0}} \sup_{\underline{v}\in\mathbb{R}^{n_V},\underline{v}\neq\underline{0}} \frac{\underline{v}^T B^T \underline{q}}{\|\underline{v}\|_2 \left\|\underline{q}\right\|_2} \geq \lambda_{\min}^{1/2}\left(BB^T\right) > 0.$$

Altogether, under the assumption that

- A is non-singular, i.e., $A : V \to V'$ is an isomorphism,
- (3.9) is satisfied,

the system matrix (3.6) is non-singular.

The result presented here just states that the given problem has a unique solution because (3.9) is satisfied. In the finite element theory it turns out that there is another important aspect to study, namely the dependency of β on the dimension of the finite element spaces. To obtain optimal orders of convergence, β has to be independent of the dimension, e.g., compare Remark 4.29. This aspect can also be taken into account in the matrix-vector formulation of linear saddle point problems, see Sect. 3.6.6. Then, one has to solve a generalized eigenvalue problem, see (3.150).

It turns out that one gets similar conditions in the general case, see Lemma 3.12 and Theorem 3.18. Whether or not these conditions are satisfied depends finally on the spaces V and Q. □

Remark 3.8 (A Manifold and a Subspace in V) A manifold of V will be defined that contains all elements which fulfill the second equation of (3.5)

$$V(r) = \{v \in V \ : \ Bv = r\}, \quad V_0 := V(0) = \ker(B).$$

The manifold V_0 is even a subspace of V. From Hilbert space theory, it follows that there is an orthogonal decomposition, with respect to the inner product of V,

$$V = V_0^{\perp} \oplus V_0,$$

where $V_0^{\perp}$ is the orthogonal complement of V_0. □

Lemma 3.9 (Properties of V_0 and $V_0^{\perp}$) *The spaces V_0 and $V_0^{\perp}$ are closed subspaces of V.*

Proof First, the closeness of V_0 will be proved. Let $\{v_n\}_{n=1}^{\infty}$ be an arbitrary Cauchy sequence with $v_n \in V_0$ for all n. Since V is complete, there exists a $v \in V$ with $\lim_{n\to\infty} v_n = v$. One has to show that $v \in V_0$. By the continuity of the linear operator B, it follows that

$$Bv = B\left(\lim_{n\to\infty} v_n\right) = \lim_{n\to\infty} (Bv_n) = \lim_{n\to\infty} 0 = 0.$$

Hence $v \in V_0$ and V_0 is closed.

The closeness of $V_0^{\perp}$ follows from the fact that the orthogonal complement of every subspace is closed, see Lemma A.17. ■

Remark 3.10 (Functionals Vanishing on V_0) A subset of V' is defined for the following analysis:

$$\tilde{V}' = \{\phi \in V' \ : \ \langle \phi, v\rangle_{V',V} = 0 \ \ \forall \ v \in V_0\} \subset V'. \tag{3.11}$$

This subset, which is even a closed subspace of V', contains all linear functionals on V that vanish for all $v \in V_0 = \ker(B)$. □

Remark 3.11 (Reduction of the System to a Single Equation in a Subspace) In the next step, the following problem is associated with Problems (3.4) and (3.5): Find $u \in V(r)$ such that

$$a(u, v) = \langle f, v\rangle_{V',V} \quad \forall \ v \in V_0. \tag{3.12}$$

Clearly, if $(u,p) \in V \times Q$ is a solution of (3.4) or (3.5), then $u \in V(r)$. In addition, one obtains

$$\left\langle B'p, v\right\rangle_{V',V} = \langle Bv, p\rangle_{Q',Q} = b(v,p) = 0 \quad \forall\ v \in V_0. \tag{3.13}$$

Since the first equation of (3.4) holds for all $v \in V$, it holds in particular for all $v \in V_0$. With (3.13) it follows that u is a solution of (3.12).

The aim of the analysis consists now in finding conditions to ensure that the converse of this statement holds: if $u \in V(r)$ is a solution of (3.12), one can find a unique $p \in Q$ such that (u,p) is the unique solution of (3.4) or (3.5), respectively.

□

Lemma 3.12 (The Inf-Sup Condition) *The three following properties are equivalent:*

i) There exists a constant $\beta_{\text{is}} > 0$ *such that*

$$\inf_{q\in Q, q\neq 0} \sup_{v\in V, v\neq 0} \frac{b(v,q)}{\|v\|_V \|q\|_Q} \geq \beta_{\text{is}}. \tag{3.14}$$

ii) The operator B' *is an isomorphism from* Q *onto* $\tilde{V}'$ *and*

$$\left\|B'q\right\|_{V'} \geq \beta_{\text{is}} \|q\|_Q \quad \forall\ q \in Q. \tag{3.15}$$

iii) The operator B *is an isomorphism from* $V_0^{\perp}$ *onto* Q' *and*

$$\|Bv\|_{Q'} \geq \beta_{\text{is}} \|v\|_V \quad \forall\ v \in V_0^{\perp}. \tag{3.16}$$

Proof The proof follows Girault and Raviart (1986).

- *i) and ii) are equivalent.*

ii) $\Longrightarrow$ *i).* From the definition of the norm of a linear functional and the definition of the adjoint operator, it follows that

$$\left\|B'q\right\|_{V'} = \sup_{v\in V, v\neq 0} \frac{\langle B'q, v\rangle_{V',V}}{\|v\|_V} = \sup_{v\in V, v\neq 0} \frac{\langle Bv, q\rangle_{Q',Q}}{\|v\|_V} = \sup_{v\in V, v\neq 0} \frac{b(v,q)}{\|v\|_V}. \tag{3.17}$$

Hence, (3.15) gives

$$\sup_{v\in V, v\neq 0} \frac{b(v,q)}{\|v\|_V} \geq \beta_{\text{is}} \|q\|_Q \quad \forall\ q \in Q.$$

Dividing by $\|q\|_Q$ and taking the infimum with respect to q on both sides of this inequality shows that ii) implies i).

i) $\Longrightarrow$ *ii).* The inequality (3.15) follows from (3.14) and (3.17). It remains to prove that B' is an isomorphism from Q onto $\tilde{V}'$. First, it will be shown that B' is an isomorphism from Q onto range(B') with a continuous inverse, see also Theorem A.70. By the definition of the range, B' is surjective on range(B'). If it would be not injective then there would be $q_1 \neq q_2 \in Q$ such that $B'q_1 = B'q_2$ or equivalently $B'(q_1 - q_2) = 0$. In this case, it follows that

$$\langle B'(q_1 - q_2), v\rangle_{V',V} = b(v, q_1 - q_2) = 0$$

for all $v \in V$ and $\|q_1 - q_2\|_Q > 0$. Then, (3.14) cannot hold, in contrast to the assumption. Hence, B' is an isomorphism from Q onto range(B').

Moreover, the inverse operator is continuous (or bounded) if and only if range(B') is a closed subspace of V'. By assumption, it is $B' \in \mathcal{L}(Q, V')$, hence B' is bounded. Thus, one can apply the Closed Range Theorem of Banach, see Theorem A.71 iv), to B' and one obtains immediately

$$\text{range}\left(B'\right) = \tilde{V}'.$$

Since $\tilde{V}'$ is a closed subspace of V', see Remark 3.10, the inverse operator is continuous. Alternatively, the closedness of range(B') follows from Theorem A.71 iii).

Altogether, B' is an isomorphism from Q onto $\tilde{V}'$.

- *ii) and iii) are equivalent.*

A standard property of an operator B and its adjoint operator B' is that range$(B') = (\text{domain}(B))'$. Hence, one has to show that the range $\tilde{V}'$ of B' can be identified with the dual space $\left(V_0^\perp\right)'$ of the domain of B.

An isomorphism between $\tilde{V}'$ and $\left(V_0^\perp\right)'$ will be constructed. The space $V_0^\perp$ is defined in Remark 3.8. Let $v \in V$ and denote the projection of v onto $V_0^\perp$ by $v^\perp$. Then, with $\phi \in \left(V_0^\perp\right)'$, one associates the element $\psi \in V'$ defined by

$$\langle \psi, v\rangle_{V',V} = \left\langle \phi, v^\perp\right\rangle_{V',V} \quad \forall\, v \in V. \tag{3.18}$$

In particular, one obtains for $v \in V_0$ that $v^\perp = 0$ and

$$\langle \psi, v\rangle_{V',V} = \left\langle \phi, v^\perp\right\rangle_{V',V} = \langle \phi, 0\rangle_{V',V} = 0 \quad \forall\, v \in V_0.$$

Hence, $\psi \in \tilde{V}'$.

The linear continuous mapping $\left(V_0^\perp\right)' \to \tilde{V}'$, $\phi \mapsto \psi$ is injective since for $\phi_0, \phi_1 \in \left(V_0^\perp\right)'$ with $\phi_0 \neq \phi_1$ and

$$\langle \psi, v\rangle_{V',V} = \left\langle \phi_0, v^\perp\right\rangle_{V',V} = \left\langle \phi_1, v^\perp\right\rangle_{V',V} \quad \forall\, v \in V,$$

it follows that

$$\left\langle \phi_0 - \phi_1, v^{\perp} \right\rangle_{V',V} = 0 \quad \forall\ v \in V.$$

Since all elements of $V_0^{\perp}$ are arguments of the linear functional $\phi_0 - \phi_1$, it follows that $\phi_0 - \phi_1$ is the null element of $\left(V_0^{\perp}\right)'$.

The mapping is also surjective. Consider an arbitrary element $\psi \in \tilde{V}'$ and an arbitrary element $v \in V$. Using the decomposition $v = v_0 + v^{\perp}$, $v_0 \in V_0$, $v^{\perp} \in V_0^{\perp}$, gives

$$\langle \psi, v \rangle_{V',V} = \langle \psi, v_0 \rangle_{V',V} + \left\langle \psi, v^{\perp} \right\rangle_{V',V} = \left\langle \psi, v^{\perp} \right\rangle_{V',V} \quad \forall\ v \in V.$$

Comparing this equation with (3.18) shows that the restriction of ψ from V to $V_0^{\perp}$ defines the inverse image of ψ. Thus, surjectivity is proved.

Altogether, the mapping is bijective and the spaces $\left(V_0^{\perp}\right)'$ and $\tilde{V}'$ can be identified. In essence, each functional from $\tilde{V}'$ is mapped to the functional that corresponds to its restriction (of its domain) from V on $V_0^{\perp}$ and there are no two functionals that both vanish on V_0 and give the same result in $V_0^{\perp}$.

The equivalence of the range of B' and the dual space of the domain of B proves the statements concerning the isomorphisms.

For proving the equivalence of the inequalities (3.15) and (3.16), first

$$\sup_{v \in V, v \neq 0} \frac{\langle Bv, q \rangle_{Q',Q}}{\|v\|_V} = \sup_{v \in V_0^{\perp}, v \neq 0} \frac{\langle Bv, q \rangle_{Q',Q}}{\|v\|_V} \tag{3.19}$$

will be shown. That the left-hand side is larger or equal than the right-hand side follows from the fact that the supremum is taken in a larger set. Using the orthogonal decomposition $v = v_0 + v^{\perp}$, $v_0 \in V_0$, $v^{\perp} \in V_0^{\perp}$, yields

$$\begin{aligned}
\sup_{v \in V, v \neq 0} \frac{\langle Bv, q \rangle_{Q',Q}}{\|v\|_V} &= \sup_{v \in V, v \neq 0} \frac{\left\langle Bv_0 + Bv^{\perp}, q \right\rangle_{Q',Q}}{\left(\|v_0\|_V^2 + \left\|v^{\perp}\right\|_V^2\right)^{1/2}} \\
&= \sup_{v \in V, v \neq 0} \frac{\left\langle Bv^{\perp}, q \right\rangle_{Q',Q}}{\left(\|v_0\|_V^2 + \left\|v^{\perp}\right\|_V^2\right)^{1/2}} \\
&\le \sup_{v \in V, v \neq 0} \frac{\left\langle Bv^{\perp}, q \right\rangle_{Q',Q}}{\left\|v^{\perp}\right\|_V} \\
&= \sup_{v \in V_0^{\perp}, v \neq 0} \frac{\langle Bv, q \rangle_{Q',Q}}{\|v\|_V},
\end{aligned}$$

since reducing the denominator can at most increase the quotient. Combining both inequalities proves (3.19). One gets from (3.15), the definition of the norm in V', the definition of the adjoint operator, (3.19), and the definition of the norm in Q'

$$
\begin{aligned}
\beta_{\mathrm{is}}^{-1} &\geq \sup_{q \in Q, q \neq 0} \frac{\|q\|_Q}{\|B'q\|_{V'}} = \sup_{q \in Q, q \neq 0} \sup_{v \in V, v \neq 0} \frac{\|q\|_Q \, \|v\|_V}{\langle B'q, v \rangle_{V',V}} \\
&= \sup_{q \in Q, q \neq 0} \sup_{v \in V, v \neq 0} \frac{\|q\|_Q \, \|v\|_V}{\langle Bv, q \rangle_{Q',Q}} = \sup_{q \in Q, q \neq 0} \sup_{v \in V_0^{\perp}, v \neq 0} \frac{\|q\|_Q \, \|v\|_V}{\langle Bv, q \rangle_{Q',Q}} \\
&= \sup_{v \in V_0^{\perp}, v \neq 0} \sup_{q \in Q, q \neq 0} \frac{\|q\|_Q \, \|v\|_V}{\langle Bv, q \rangle_{Q',Q}} = \sup_{v \in V_0^{\perp}, v \neq 0} \frac{\|v\|_V}{\|Bv\|_{Q'}}.
\end{aligned}
$$

From this estimate, (3.16) follows. Analogously, one can derive (3.15) from (3.16). Thus, both estimates are equivalent. ■

Theorem 3.13 (Inf-Sup Condition for the Complete Bilinear Form, Babuška's Inf-Sup Condition) *Let V_1 and V_2 be two Hilbert spaces with inner products $(\cdot,\cdot)_{V_1}$ and $(\cdot,\cdot)_{V_2}$. Let $\mathcal{A}(\cdot,\cdot)$ be a bilinear form on $V_1 \times V_2$ such that*

$$
\begin{aligned}
|\mathcal{A}(v_1, v_2)| &\leq C_1 \, \|v_1\|_{V_1} \, \|v_2\|_{V_2}, \quad \forall \ v_1 \in V_1, v_2 \in V_2, \\
\sup_{v_1 \in V_1, \|v_1\|_{V_1} \neq 0} \frac{\mathcal{A}(v_1, v_2)}{\|v_1\|_{V_1}} &\geq C_2 \, \|v_2\|_{V_2} \quad \forall \ v_2 \in V_2, \\
\sup_{v_2 \in V_2, \|v_2\|_{V_2} \neq 0} \frac{\mathcal{A}(v_1, v_2)}{\|v_2\|_{V_2}} &\geq C_3 \, \|v_1\|_{V_1} \quad \forall \ v_1 \in V_1,
\end{aligned}
$$

with $C_1 < \infty$ and $C_2, C_3 > 0$. Let $f \in V_2'$, then there exists exactly one element $u \in V_1$ such that

$$
\mathcal{A}(u, v_2) = \langle f, v_2 \rangle_{V_2', V_2} \quad \forall \ v_2 \in V_2
$$

and

$$
\|u\|_{V_1} \leq \frac{1}{C_2} \, \|f\|_{V_2'}.
$$

Proof For the proof, it is referred to Babuška (1971). ■

Remark 3.14 (Inf-Sup Conditions, Babuška–Brezzi Condition) Condition (3.14) is called inf-sup condition. It was introduced in this form in Brezzi (1974). In Babuška (1971), the inf-sup condition presented in Theorem 3.13 was proved. Choosing in this theorem $V_1 = V_2 = V \times Q$ and setting the bilinear form to be

$$
\mathcal{A}(v_1, v_2) = \mathcal{A}((v + w), (q + r)) = a(v, w) + b(w, q) + b(v, r)
$$

leads to the inf-sup condition

$$\sup_{\substack{(v,q)\in V\times Q\\ \|v\|_V+\|q\|_Q\neq 0}} \frac{a(v,w)+b(w,q)+b(v,r)}{\|v\|_V+\|q\|_Q} \ge \beta_{\mathrm{is,Bab}}\left(\|w\|_V+\|r\|_Q\right)$$

for all $(w,r) \in V \times Q$. If $a(\cdot,\cdot)$ is symmetric, then both inf-sup conditions in Theorem 3.13 are identical. The inf-sup condition from Babuška (1971) was applied to the analysis of finite element problems with Lagrangian multipliers in Babuška (7273). Note that the pressure in the Navier–Stokes equations can be interpreted as a Lagrangian multiplier associated with the imposition of the divergence-free constraint for the velocity.

Nevertheless, it became common to call (3.14) Babuška–Brezzi condition. The relation between the inf-sup conditions from Babuška (1971, 7273) and Brezzi (1974) in the context of finite element methods is discussed briefly in Remark 3.54.

Sometimes, the inf-sup condition (3.14) is even called Ladyzhenskaya–Babuška–Brezzi condition or LBB condition. It is mentioned in the literature, e.g., in Gunzburger (2002), that in the book Ladyzhenskaya (1969) the property (3.16) is proved. The proof of the uniqueness of the pressure in Ladyzhenskaya (1969, Chap. 2.1) uses an argument that is based on the uniqueness of the Helmholtz decomposition of vector fields in $L^2(\Omega)$, see Sect. 3.7. □

Remark 3.15 (On the Inf-Sup Condition and the Coercivity) For a bounded bilinear form $c : V \times V \to \mathbb{R}$, the theorem of Lax–Milgram, see Theorem B.4, states that the coercivity of $c(\cdot,\cdot)$ is a sufficient condition for the existence of a unique solution of a corresponding problem. The coercivity can be written in the form that there is a constant $m > 0$

$$m\,\|w\|_V^2 \le c(w,w) \quad \forall\, w \in V \quad \Longleftrightarrow \quad m\,\|w\|_V \le \frac{c(w,w)}{\|w\|_V} \quad \forall\, w \in V.$$

From the last formulation, it follows that also

$$m\,\|w\|_V \le \sup_{v\in V, v\neq 0} \frac{c(v,w)}{\|v\|_V} \quad \forall\, w \in V, \tag{3.20}$$

holds. Observe that in this more general condition, the spaces in the bilinear form are allowed to be different.

Assume now that $u \in V(r)$ is the unique solution of (3.12). Then, the first equation of (3.4) can be written in the form

$$b(v,p) = \langle f, v\rangle_{V',V} - a(u,v) \quad \forall\, v \in V.$$

The concept of coercivity cannot be applied for the analysis of this equation. But note that the inf-sup condition (3.14) can be written in the form

$$\beta_{\mathrm{is}} \|q\|_Q \leq \sup_{v\in V, v\neq 0} \frac{b(v,q)}{\|v\|_V} \quad \forall\, q \in Q,$$

which is the same form as (3.20).

Altogether, the inf-sup condition can be viewed as a generalization of the coercivity condition to bilinear forms where the arguments are from different spaces. □

Remark 3.16 (The Space $V(r)$ is Not Empty) From the inf-sup condition, one gets that V(r) is not empty. Let $r \in Q'$, then it follows from Lemma 3.12 iii) that there is a $v \in V_0^{\perp}$ such that $Bv = r$. □

Remark 3.17 (An Imbedding Operator) Before stating the results concerning the well-posedness of Problem (3.5), see Definition 3.5, the linear continuous operator $E_0 \in \mathcal{L}\left(V', V_0'\right)$ given by

$$\langle E_0\phi, v\rangle_{V',V} = \langle \phi, v\rangle_{V',V} \quad \forall\, \phi \in V',\ \forall\, v \in V_0, \tag{3.21}$$

is introduced. This operator is an imbedding operator. Since $V_0 \subset V$, it is $V_0' \supset V'$ and $E_0\phi$ is the restriction of the functional ϕ from V onto V_0. The subspace V_0 is equipped with the same norm as V, hence one obtains

$$\begin{aligned}\|E_0\phi\|_{V_0'} &= \sup_{v\in V_0, v\neq 0} \frac{\langle E_0\phi, v\rangle_{V',V}}{\|v\|_V} = \sup_{v\in V_0, v\neq 0} \frac{\langle \phi, v\rangle_{V',V}}{\|v\|_V} \\ &\leq \sup_{v\in V, v\neq 0} \frac{\langle \phi, v\rangle_{V',V}}{\|v\|_V} = \|\phi\|_{V'}\,. \end{aligned} \tag{3.22}$$

□

Theorem 3.18 (Well-posedness of Problem (3.5)**)** *Problem* (3.5) *is well-posed if and only if the following two conditions hold:*

i) The operator $E_0 \circ A$ is an isomorphism from V_0 onto V_0'.
ii) The bilinear form $b(\cdot,\cdot)$ satisfies the inf-sup condition (3.14)*.*

Proof • *The conditions i) and ii) are sufficient.*

Assume that conditions i) and ii) are satisfied. In the first step, it will be proved that Problem (3.12) possesses a unique solution $u \in V(r)$. Then, the second step proves the existence of a unique solution $(u,p) \in V \times Q$ of Problem (3.5).

Unique Solution $u \in V(r)$ of Problem (3.12) It follows from Lemma 3.12 iii) that there is a unique $u_0 \in V_0^{\perp}$ such that

$$Bu_0 = r, \quad \|u_0\|_V \leq \frac{1}{\beta_{\mathrm{is}}} \|r\|_{Q'}\,. \tag{3.23}$$

Therefore, Problem (3.12) can be stated equivalently in the following way: Find $w = u - u_0 \in V_0$ such that

$$a(w, v) = \langle f, v\rangle_{V',V} - a(u_0, v) \quad \forall\ v \in V_0.$$

Now, the ansatz and the test space are the same. Writing the equation in the form

$$\langle Aw, v\rangle_{V',V} = \langle f - Au_0, v\rangle_{V',V} \quad \forall\ v \in V_0, \tag{3.24}$$

then it follows from (3.21) that w satisfies the operator equation

$$(E_0 \circ A)\, w = E_0 \circ (f - Au_0) \quad \text{in } V_0'. \tag{3.25}$$

By assumption i), $E_0 \circ A$ is an isomorphism from V_0 onto V_0', such there is a unique w satisfying this equation. Hence, Problem (3.12) has the unique solution $u = u_0 + w \in V(r)$.

Unique Solution $(u, p) \in V \times Q$ *of Problem* (3.5) From the boundedness of A and E_0, it follows that $E_0 \circ A$ is bounded. Therefore, the inverse $(E_0 \circ A)^{-1}$ is a continuous bounded operator, see Theorem A.70. From (3.25), this property, and (3.22), it follows that

$$\begin{aligned}\|w\|_V &= \left\|(E_0 \circ A)^{-1}\, E_0 \circ (f - Au_0)\right\|_V \le C\, \|E_0 \circ (f - Au_0)\|_{V_0'} \\ &\le C\, \|f - Au_0\|_{V'}\,.\end{aligned}$$

The application of the triangle inequality, the boundedness of the operator A, and twice (3.23) yields

$$\begin{aligned}\|u\|_V \le \|u_0\|_V + \|w\|_V &\le \frac{1}{\beta_{\text{is}}}\, \|r\|_{Q'} + C\, \|f - Au_0\|_{V'} \\ &\le C\left(\|r\|_{Q'} + \|f\|_{V'} + \|u_0\|_V\right) \le C\left(\|r\|_{Q'} + \|f\|_{V'}\right).\end{aligned} \tag{3.26}$$

Using the decomposition of u and (3.24) yields $f - Au = f - Au_0 - Aw \in \tilde{V}'$. Thus, according to Lemma 3.12 ii) there exists a unique $p \in Q$ such that

$$B'p = f - Au.$$

Using Lemma 3.12 ii), the triangle inequality, the boundedness of A, and (3.26), one obtains

$$\begin{aligned}\|p\|_Q &\le \frac{1}{\beta_{\text{is}}}\, \|f - Au\|_{V'} \le \frac{1}{\beta_{\text{is}}} \left(\|f\|_{V'} + \|Au\|_{V'}\right) \\ &\le \frac{1}{\beta_{\text{is}}} \left(\|f\|_{V'} + C\, \|u\|_V\right) \le C\left(\|f\|_{V'} + \|r\|_{Q'}\right).\end{aligned} \tag{3.27}$$

Hence, Problem (3.5) has a unique solution (u, p) and the norm of the solution is bounded by the norm of the data of the problem, which is stability. Or in other words, from (3.26) and (3.27), it follows that the mapping from the right-hand side to the solution $(f, r) \mapsto (u, p)$ is bounded. Thus, this mapping is continuous from $V' \times Q'$ onto $V \times Q$. This property means, together with the already proved uniqueness of u and p, that Φ is an isomorphism from $V \times Q$ onto $V' \times Q'$.

- *The conditions i) and ii) are necessary.*

Assume now that Φ is an isomorphism from $V \times Q$ onto $V' \times Q'$. It will be proved in the first step that then the inf-sup condition (3.14) holds, i.e., condition ii) of the theorem. The second step proves that also condition i) of the theorem is satisfied.

Condition ii) of the Theorem Holds Consider the restriction of B to $V_0^\perp$ and denote it by $B^\perp$. Every $u \in V$ can be decomposed uniquely into $u = u_0 + \tilde{u}$, $u_0 \in V_0$, $\tilde{u} \in V_0^\perp$. Then, one obtains, using the definition of V_0,

$$Bu = Bu_0 + B\tilde{u} = B\tilde{u} = B^\perp \tilde{u}. \tag{3.28}$$

Since Φ is an isomorphism, it is range $(B) = Q'$ and because of (3.28), one gets range $\left(B^\perp\right) = Q'$. Hence, $B^\perp$ is surjective. Next, one has to show that $B^\perp$ is injective. Let $u_0, u_1 \in V_0^\perp$ with $B^\perp u_0 = B^\perp u_1 = \psi$. Hence $B^\perp(u_0 - u_1) = 0$ from what follows that $u_0 - u_1 \in V_0$. Because $u_0, u_1 \in V_0^\perp$ also each linear combination is element of $V_0^\perp$, in particular $u_0 - u_1 \in V_0^\perp$. The only element which is in V_0 and $V_0^\perp$ is the zero element such that $u_0 = u_1$ follows and $B^\perp$ is injective. Altogether, $B^\perp$ is an isomorphism between $V_0^\perp$ and Q'. Consequently, the inverse map $(B^\perp)^{-1}$ is also an isomorphism. Theorem A.70 gives that it is a bounded operator, i.e., for all $\psi \in Q'$ there is a constant C such that

$$\left\| (B^\perp)^{-1} \psi \right\|_V \le C \left\| \psi \right\|_{Q'}.$$

Choosing $\psi = Bv$, one observes that this inequality is equivalent to the existence of a constant β_{is} such that for all $v \in V_0^\perp$

$$\beta_{\text{is}} \left\| v \right\|_V \le \left\| Bv \right\|_{Q'}.$$

From Lemma 3.12 it follows that then condition (3.14) holds.

Condition i) of the Theorem Holds Let $\phi \in V_0'$ be arbitrary. By the Hahn–Banach Theorem, Theorem A.72, there exists at least one element $f \in V'$ such that $\phi = E_0 f$. Setting $(u, p) = \Phi^{-1}(f, 0)$, then one has $u \in V_0$, since $Bu = 0$, and

$$Au + B'p = f \quad \text{in } V'. \tag{3.29}$$

By the definition of E_0 it follows that for all $v \in V_0$

$$\left\langle E_0 \circ B'p, v \right\rangle_{V',V} = \left\langle B'p, v \right\rangle_{V',V} = \left\langle Bv, p \right\rangle_{Q',Q} = 0.$$

That means $E_0 \circ B'p = 0$ in V_0' and one gets with (3.29)

$$E_0 \circ Au = E_0 f = \phi \quad \text{in } V_0'.$$

Hence, for each $\phi \in V_0'$ there is an element $u \in V_0$ with $E_0 \circ Au = \phi$, such that $E_0 \circ A$ is surjective.

Now, the injectivity of $E_0 \circ A$ will be proved. Let $v_1 \neq v_2 \in V_0$ with

$$E_0 \circ Av_1 = E_0 \circ Av_2 = \phi \quad \Longrightarrow \quad E_0 \circ A(v_1 - v_2) = 0. \tag{3.30}$$

Let $q \in Q$ be arbitrary. Then

$$Av_1 + B'q = f_1, \quad Av_2 + B'q = f_2 \quad \in V',$$

with $f_1 \neq f_2$ since Φ is an isomorphism and $Bv_1 = Bv_2 = 0$. It follows that

$$A(v_1 - v_2) = f_1 - f_2 \quad \in V' \quad \Longleftrightarrow \quad \Phi(v_1 - v_2, 0) = (f_1 - f_2, 0). \tag{3.31}$$

By the definition (3.21) of E_0, one obtains with (3.30) and (3.31)

$$0 = \langle E_0 \circ A(v_1 - v_2), v\rangle_{V',V} = \langle A(v_1 - v_2), v\rangle_{V',V} = \langle f_1 - f_2, v\rangle_{V',V} \quad \forall\, v \in V_0.$$

With the definition of $\tilde{V}'$ it follows that $f_1 - f_2 \in \tilde{V}'$. It was already proved that the inf-sup condition holds. Thus, it follows from Lemma 3.12 ii) that there is a unique $q \in Q$ such that $B'q = f_1 - f_2$. Hence, $\Phi(0, q) = (f_1 - f_2, 0)$. Since Φ is an isomorphism, it follows with (3.31) that $(v_1 - v_2, 0) = (0, q)$, and in particular that $v_1 = v_2$.

Thus, $E_0 \circ A$ is a one-to-one linear continuous mapping from V_0 onto V_0' and therefore an isomorphism from V_0 onto V_0'. ■

Lemma 3.19 (Sufficient Condition on $a(\cdot,\cdot)$ for the Well-posedness of (3.5)) *Assume that the bilinear form $a(\cdot,\cdot)$ is V_0-elliptic, i.e., there is a constant $\alpha > 0$ such that*

$$a(v, v) \geq \alpha \|v\|_V^2 \quad \forall\, v \in V_0.$$

Then, Problem (3.5) is well-posed if and only if the bilinear form $b(\cdot,\cdot)$ satisfies the inf-sup condition (3.14).

Proof • *V_0-ellipticity and inf-sup condition $\Longrightarrow$ well-posed problem.* It will be shown that the V_0-ellipticity implies the first condition of Theorem 3.18. Let $f \in V_0'$ be arbitrary. Since $a(\cdot,\cdot)$ is V_0-elliptic, the Lax–Milgram theorem, see Theorem B.4, gives that there is a unique $u \in V_0$ such that

$$a(u, v) = \langle f, v\rangle_{V',V} \quad \forall\, v \in V_0,$$

or equivalently, that there is a unique $u \in V_0$ such that

$$E_0 \circ Au = f.$$

Since f was chosen to be arbitrary, $E_0 \circ A$ is surjective.

Consider now the injectivity of $E_0 \circ A$. Let $u \in V_0$ with $E_0 \circ Au = 0$. By the definition (3.21) of E_0, it is

$$0 = \langle E_0 \circ Au, v\rangle_{V',V} = \langle Au, v\rangle_{V',V} \quad \forall\ v \in V_0,$$

and in particular

$$0 = \langle Au, u\rangle_{V',V} = a(u, u).$$

From the V_0-ellipticity it follows that $u = 0$, which implies injectivity.

Altogether, $E_0 \circ A$ is an isomorphism from V_0 onto V_0' and Theorem 3.18 gives the well-posedness of Problem 3.5.

- *V_0-ellipticity and well-posed problem $\Longrightarrow$ inf-sup condition.* The satisfaction of the inf-sup condition in case Problem 3.5 is well-posed follows directly from Theorem 3.18. The V_0-ellipticity of the bilinear form $a(\cdot,\cdot)$ is a special case of the first condition of Theorem 3.18.

■

Remark 3.20 (Formulation as an Optimization Problem, Saddle Point Problem) The Problems (3.5) and (3.12) can formulated as optimization problems under certain conditions. Let $J_0 : V \to \mathbb{R}$ and $J_1 : V \times Q \to \mathbb{R}$ be two quadratic functionals defined by

$$J_0(v) = \frac{1}{2}a(v, v) - \langle f, v\rangle_{V',V}, \quad J_1(v, q) = J_0(v) + b(v, q) - \langle r, q\rangle_{Q',Q}.$$

The functional J_0 is called energy functional associated with Problem (3.12) and J_1 is the Lagrangian functional associated with Problem (3.5).

Consider the following problem: Find a saddle point $(u, p) \in V \times Q$ of the Lagrangian functional J_1 over $V \times Q$, i.e., find a pair $(u, p) \in V \times Q$ such that

$$J_1(u, q) \le J_1(u, p) \le J_1(v, p) \quad \forall\ v \in V,\ \forall\ q \in Q. \tag{3.32}$$

This form is the classical formulation of a saddle point problem. The characterization (3.32) inspired the notation saddle point problem also for Problem (3.5). □

Theorem 3.21 (Existence and Uniqueness of a Solution of (3.32)) *Assume the conditions i) and ii) of Theorem 3.18. Assume in addition that the bilinear form $a(\cdot,\cdot)$ is symmetric and positive semi-definite on V, i.e.,*

$$a(v, v) \ge 0 \quad \forall\ v \in V.$$

Then, Problem (3.32) *has a unique solution* $(u, p) \in V \times Q$ *that is precisely the solution of Problem* (3.5).

Proof It is referred to Girault and Raviart (1986, p. 62) for the proof. ■

Remark 3.22 (Some Generalizations) Generalizations of the saddle point problem considered in this section have been studied in the literature.

- In Ciarlet et al. (2003), solvability and stability conditions for a saddle point problem of the form

$$\begin{aligned} a(u, v) + b_1(v, p) &= \langle f, v\rangle_{V',V} \quad \forall\ v \in V, \\ b_2(u, q) - c(p, q) &= \langle r, q\rangle_{Q',Q} \quad \forall\ q \in Q, \end{aligned}$$

  are established.
- Abstract saddle point problems of form (3.4) can be also studied in Banach spaces, see Ern and Guermond (2004, Chap. 2.4) and the references therein.

□

3.2 Appropriate Function Spaces for Continuous Incompressible Flow Problems

Remark 3.23 (Contents) The theory of Sect. 3.1 will now be applied to characterize appropriate function spaces for weak formulations of incompressible flow problems. Theorem 3.18 and Lemma 3.19 give two conditions for the well-posedness of the linear saddle point problem. One condition concerns only the space V. It will be discussed for the individual incompressible flow models later, e.g., see Theorem 4.6 for the Stokes equations. The emphasis of this section is on the second condition, which establishes a connection between the spaces V and Q. These spaces have to satisfy the inf-sup condition (3.14). Note that the inf-sup condition guarantees the uniqueness of the pressure, e.g., see the proof of Theorem 3.18. □

Remark 3.24 (The Bilinear Form $b(\cdot,\cdot)$ *for Incompressible Flow Problems)* In the inf-sup condition (3.14), the velocity and pressure space are coupled by a bilinear form. A weak formulation of incompressible flow problems is obtained in the usual way by multiplying the momentum equation with a test function $\boldsymbol{v} \in V$ and the continuity equation with a test function $q \in Q$. Then, both equations are integrated on Ω. One obtains for the continuity equation

$$\int_\Omega (\nabla \cdot \boldsymbol{u})\, q\, d\boldsymbol{x} = (\nabla \cdot \boldsymbol{u}, q) = 0.$$

For the viscous term and the pressure term in the continuity equation, integration by parts is applied. Assuming that the functions are sufficiently smooth and that the integral on the boundary vanishes in performing the integration by parts, one gets the term

$$\int_{\Omega} \nabla p \cdot \boldsymbol{v}\, d\boldsymbol{x} = -\int_{\Omega} (\nabla \cdot \boldsymbol{v})\, p\, d\boldsymbol{x} = -(\nabla \cdot \boldsymbol{v}, p). \tag{3.33}$$

Thus, the framework of Sect. 3.1 can be used if one defines

$$b(\boldsymbol{v}, q) = -\int_{\Omega} (\nabla \cdot \boldsymbol{v})\, q\, d\boldsymbol{x} = -(\nabla \cdot \boldsymbol{v}, q) \quad \boldsymbol{v} \in V,\ q \in Q. \tag{3.34}$$

□

Remark 3.25 (Function Spaces for Velocity and Pressure for Homogeneous Dirichlet Boundary Conditions) Let Ω be a bounded and connected domain in $\mathbb{R}^d$, $d \in \{2, 3\}$, with Lipschitz boundary. To simplify the presentation, only problems with Dirichlet boundary conditions on the whole boundary will be considered. Since these are essential boundary conditions, they enter the definition of the velocity space. Define

$$V = H_0^1(\Omega) = \left\{\boldsymbol{v} \ :\ \boldsymbol{v} \in H^1(\Omega) \text{ with } \boldsymbol{v} = \boldsymbol{0} \text{ on } \Gamma\right\},$$

where the value of $\boldsymbol{v}$ on the boundary is to be understood in the sense of traces, and

$$Q = L_0^2(\Omega) = \left\{q \ :\ q \in L^2(\Omega) \text{ with } \int_{\Omega} q(\boldsymbol{x})\, d\boldsymbol{x} = 0\right\}.$$

Both spaces are Hilbert spaces. The inner product in V and the induced norm are given by

$$(\boldsymbol{v}, \boldsymbol{w}) = \int_{\Omega} (\nabla \boldsymbol{v} \cdot \nabla \boldsymbol{w})(\boldsymbol{x})\, d\boldsymbol{x}, \quad \|\boldsymbol{v}\|_V = \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}. \tag{3.35}$$

Poincaré's inequality (A.12) shows that (3.35) defines in fact an inner product and a norm in V. The inner product and the induced norm in Q are given by

$$(q, r) = \int_{\Omega} (qr)(\boldsymbol{x})\, d\boldsymbol{x}, \quad \|q\|_Q = \|q\|_{L^2(\Omega)}.$$

The dual space of V is $V' = H^{-1}(\Omega)$ and the dual of the pressure space is $Q' = Q$.

For $\boldsymbol{v} \in V$ it follows that $\nabla \boldsymbol{v} \in L^2(\Omega)$ and with estimate (3.41) proved below, one obtains that $\nabla \cdot \boldsymbol{v} \in L^2(\Omega)$. Thus, the definition of the spaces implies that all terms in (3.33) are well defined and that this equality holds. □

Remark 3.26 (Notation for Spaces of Vector-valued and Tensor-valued Functions) For simplicity of notation, spaces of vector-valued or tensor-valued are denoted with the same symbol as the corresponding space for scalar functions. This notation has to be understood in the sense that each component of the vector-valued or tensor-valued function belongs to this space. □

Remark 3.27 (Spaces for Other Boundary Conditions) In applications, Dirichlet boundary conditions are often given only on a part Γ_{diri} of the boundary Γ. These boundary conditions enter still the definition of the velocity space. Other boundary conditions will appear in the bilinear forms of the weak formulation of the equations.

If outflow boundary conditions (2.37) are prescribed on a part Γ_{donot} of the boundary, then these boundary conditions describe also the pressure at Γ_{donot}. As already discussed in Remark 2.27, in this situation there is no need to fix an additive constant of the pressure and the pressure space is given by $Q = L^2(\Omega)$.

Since the analysis depends on the actual spaces and the spaces depend on the boundary conditions, different boundary conditions might require different analytical tools and they might lead to different results. Here, only the basic case, which is also the most simple one, of prescribing homogeneous Dirichlet boundary conditions on the whole boundary Γ will be considered. □

Remark 3.28 (The Divergence Operator) The divergence operator is defined by

$$\text{div} \ : \ V \to \text{range(div)}, \quad \boldsymbol{v} \mapsto \nabla \cdot \boldsymbol{v}.$$

From (3.41) below, one gets for $\boldsymbol{v} \in V$ that $\nabla \cdot \boldsymbol{v} \in L^2(\Omega)$. Integration by parts gives

$$\int_\Omega (\nabla \cdot \boldsymbol{v})\,(\boldsymbol{x})\, d\boldsymbol{x} = 0 \quad \forall\ \boldsymbol{v} \in V,$$

such that the integral mean value is zero and hence $\text{range(div)} \subseteq Q = Q'$ can be concluded. In Lemma 3.43 it will be shown that even equality holds: $\text{range(div)} = Q'$. It follows from Lemma 3.12 iii), that this condition necessarily holds if the inf-sup condition is satisfied.

Altogether, the operator $B \in \mathcal{L}(V, Q')$ from Sect. 3.1 can be characterized in incompressible flow problems as the negative divergence operator. □

Remark 3.29 (The Gradient Operator) The gradient operator will be defined on Q

$$\text{grad} \ : \ Q \to \text{range(grad)}, \quad q \mapsto \nabla q.$$

Since the gradient of a function from $L^2(\Omega)$ is in $H^{-1}(\Omega)$, one obtains $\text{range(grad)} \subset V'$. The range of grad will be characterized more precisely in Lemma 3.41, accordingly to the condition from Lemma 3.12, ii).

Integration by parts gives

$$\begin{aligned}\langle -\text{div}(\boldsymbol{v}), q\rangle_{Q',Q} &= -\int_\Omega (\nabla \cdot \boldsymbol{v}) q\, d\boldsymbol{x} = \int_\Omega \nabla q \cdot \boldsymbol{v}\, d\boldsymbol{x} \\ &= \langle \text{grad}(q), \boldsymbol{v}\rangle_{V',V} \quad \forall\ \boldsymbol{v} \in V, q \in Q. \end{aligned} \tag{3.36}$$

From this identity it follows that −div and grad are dual operators and grad represents the operator $B' \in \mathcal{L}(Q, V')$ from Sect. 3.1. □

Definition 3.30 (Distributional and Weak Divergence) For a vector field $\boldsymbol{v} \in L^1(\Omega)$, the mapping

$$C_0^\infty(\Omega) \to \mathbb{R}, \quad \psi \mapsto \int_\Omega \nabla\psi \cdot \boldsymbol{v}\, d\boldsymbol{x}$$

is called the distributional divergence of $\boldsymbol{v}$.

If for a vector field $\boldsymbol{v} \in L^p(\Omega)$ with $p \geq 1$ there exists a function $\theta \in L^1_{\text{loc}}(\Omega)$ such that

$$-\int_\Omega \nabla\psi \cdot \boldsymbol{v}\, d\boldsymbol{x} = \int_\Omega \psi\theta\, d\boldsymbol{x} \quad \forall\ \psi \in C_0^\infty(\Omega),$$

then the function θ is called the weak divergence of $\boldsymbol{v}$. □

Remark 3.31 (A Space of Functions with Weak Divergence) For incompressible flow problems, the space of vector fields in $L^2(\Omega)$ where the divergence belongs also to $L^2(\Omega)$

$$H(\text{div}, \Omega) = \left\{\boldsymbol{v} \in L^2(\Omega)\ :\ \nabla \cdot \boldsymbol{v} \in L^2(\Omega)\right\} \tag{3.37}$$

is important. The space $H(\text{div}, \Omega)$ is a Hilbert space with the inner product and the induced norm, respectively,

$$\begin{aligned} (\boldsymbol{v}, \boldsymbol{w})_{H(\text{div},\Omega)} &= (\boldsymbol{v}, \boldsymbol{w}) + (\nabla \cdot \boldsymbol{v}, \nabla \cdot \boldsymbol{w}), \\ \|\boldsymbol{v}\|_{H(\text{div},\Omega)} &= \left(\|\boldsymbol{v}\|^2_{L^2(\Omega)} + \|\nabla \cdot \boldsymbol{v}\|^2_{L^2(\Omega)}\right)^{1/2}. \end{aligned}$$

□

Definition 3.32 (Divergence-free Vector Field) In view of Definition 3.30, a vector field $\boldsymbol{v} \in L^p(\Omega)$, $p \geq 1$, is called to be weakly divergence-free if

$$\int_\Omega \nabla\psi \cdot \boldsymbol{v}\, d\boldsymbol{x} = 0 \quad \forall\ \psi \in C_0^\infty(\Omega).$$

□

Remark 3.33 (Spaces of Weakly Divergence-free Functions) It became clear in Sect. 3.1, Remark 3.8, that the kernel of the operator B is of importance. This kernel is the space of weakly divergence-free functions in V

$$V_0 = V_{\text{div}} = \{\boldsymbol{v} \in V\ :\ (\nabla \cdot \boldsymbol{v}, q) = 0\ \forall\ q \in Q\}. \tag{3.38}$$

Thus, the divergence of the functions from V_{div} vanishes in the sense of $L^2(\Omega)$, i.e., it is $(\nabla \cdot \boldsymbol{v})(\boldsymbol{x}) = 0$ almost everywhere in Ω.

Another space of divergence-free functions is defined by

$$H_{\mathrm{div}}(\Omega) = \{\boldsymbol{v} \in H(\mathrm{div}, \Omega) \ : \ \nabla \cdot \boldsymbol{v} = 0 \text{ and } \boldsymbol{v} \cdot \boldsymbol{n} = 0 \text{ on } \Gamma \text{ in the sense of traces}\}. \tag{3.39}$$

The regularity requirement for functions from $H_{\mathrm{div}}(\Omega)$ is weaker than for functions from V_{div}.

For bounded domains with Lipschitz boundary, it can be shown that $H_{\mathrm{div}}(\Omega)$ is the closure of $C_{0,\mathrm{div}}^{\infty}(\Omega)$, see (A.7), in the norm $\|\cdot\|_{L^2(\Omega)}$, e.g., see Constantin and Foias (1988, Proposition 1.8) or Sohr (2001, Chap. II, Lemma 2.5.3). □

Lemma 3.34 (Estimating the $L^2(\Omega)$ Norm of the Divergence by the $L^2(\Omega)$ Norm of the Gradient for Functions from $H^1(\Omega)$) *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, and let $\boldsymbol{v} \in H^1(\Omega)$, then it holds*

$$\|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)} \le \sqrt{d}\, \|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \quad \forall\ \boldsymbol{v} \in H^1(\Omega). \tag{3.40}$$

This estimate is sharp.

Proof Let $\boldsymbol{v} = (v_1, \ldots, v_d)^T$, then the Cauchy–Schwarz inequality for sums (A.2) and the extension of a sum by non-negative terms gives

$$\begin{aligned}
\|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)}^2 &= \int_\Omega \left(\sum_{i=1}^d \frac{\partial v_i}{\partial x_i}(\boldsymbol{x}) \right)^2 d\boldsymbol{x} \le \int_\Omega \left(\sum_{i=1}^d 1 \right) \left(\sum_{i=1}^d \left(\frac{\partial v_i}{\partial x_i} \right)^2 \right) (\boldsymbol{x})\, d\boldsymbol{x} \\
&= d \int_\Omega \sum_{i=1}^d \left(\frac{\partial v_i}{\partial x_i} \right)^2 (\boldsymbol{x})\, d\boldsymbol{x} \le d \int_\Omega \sum_{i,j=1}^d \left(\frac{\partial v_i}{\partial x_j} \right)^2 (\boldsymbol{x})\, d\boldsymbol{x} \\
&= d\, \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}^2,
\end{aligned}$$

which proves (3.40).

Let $\boldsymbol{v}(\boldsymbol{x}) = (x, y, z)^T$ in an arbitrary domain Ω. Then, one finds

$$\|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)}^2 = \int_\Omega (\nabla \cdot \boldsymbol{v})^2(\boldsymbol{x})\, d\boldsymbol{x} = \int_\Omega 3^2\, d\boldsymbol{x} = 9\, |\Omega|$$

and

$$\|\nabla \boldsymbol{v}\|_{L^2(\Omega)}^2 = \int_\Omega \nabla \boldsymbol{v} : \nabla \boldsymbol{v}\, d\boldsymbol{x} = \int_\Omega 3\, d\boldsymbol{x} = 3\, |\Omega|.$$

Hence, the factor $\sqrt{d}$ in (3.40) cannot be improved. ■

Remark 3.35 (Improvement of (3.40) for Functions from $H_0^1(\Omega)$) It will be shown in Lemma 3.179 that for functions from $V = H_0^1(\Omega)$ it holds

$$\|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)} \le \|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \quad \forall\ \boldsymbol{v} \in H_0^1(\Omega). \tag{3.41}$$

Since the proof of this estimate uses some tools which are introduced in Sect. 3.7, it is postponed to this section. □

Remark 3.36 (Estimating the Norm of the Deformation Tensor by the Norm of the Gradient) Using the triangle inequality and that the norm of a tensor is defined component-by-component, one obtains readily

$$\begin{aligned}\|\mathbb{D}(\boldsymbol{u})\|_{L^2(\Omega)} &= \left\|\frac{\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T}{2}\right\|_{L^2(\Omega)} \\ &\le \frac{1}{2}\|\nabla \boldsymbol{u}\|_{L^2(\Omega)} + \frac{1}{2}\left\|(\nabla \boldsymbol{u})^T\right\|_{L^2(\Omega)} = \|\nabla \boldsymbol{u}\|_{L^2(\Omega)}.\end{aligned}$$

Thus, the norm of the symmetric part of the gradient can be estimated by the norm of the gradient. There is also an estimate in the other direction, which is called Korn's inequality. □

Lemma 3.37 (Korn's Inequality in V) *For all $\boldsymbol{v} \in V$ it holds*

$$2\|\mathbb{D}(\boldsymbol{v})\|_{L^2(\Omega)}^2 = \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}^2 + \|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)}^2. \tag{3.42}$$

Consequently, it is

$$\|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \le \sqrt{2}\,\|\mathbb{D}(\boldsymbol{v})\|_{L^2(\Omega)}, \tag{3.43}$$

which is called Korn's inequality.

Proof Let $\boldsymbol{v} \in V$. Integration by parts gives

$$(\nabla \cdot (\nabla \boldsymbol{v}), \boldsymbol{v}) = -(\nabla \boldsymbol{v}, \nabla \boldsymbol{v}) = -\|\nabla \boldsymbol{v}\|_{L^2(\Omega)}^2.$$

Using (2.26) and again integration by parts yields

$$\left(\nabla \cdot \left(\nabla \boldsymbol{v}^T\right), \boldsymbol{v}\right) = (\nabla(\nabla \cdot \boldsymbol{v}), \boldsymbol{v}) = -\|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)}^2.$$

Applying the definition of the deformation tensor leads to

$$-2(\nabla \cdot (\mathbb{D}(\boldsymbol{v})), \boldsymbol{v}) = -(\nabla \cdot (\nabla \boldsymbol{v}), \boldsymbol{v}) - \left(\nabla \cdot \left(\nabla \boldsymbol{v}^T\right), \boldsymbol{v}\right) = \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}^2 + \|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)}^2. \tag{3.44}$$

On the other hand, one finds with integration by parts and the symmetry of the deformation tensor

$$
\begin{aligned}
-2\left(\nabla\cdot\left(\mathbb{D}\left(\boldsymbol{v}\right)\right),\boldsymbol{v}\right) &= 2\left(\mathbb{D}\left(\boldsymbol{v}\right),\nabla\boldsymbol{v}\right) = \left(\mathbb{D}\left(\boldsymbol{v}\right),\nabla\boldsymbol{v}\right) + \left(\left(\mathbb{D}\left(\boldsymbol{v}\right)\right)^T,\left(\nabla\boldsymbol{v}\right)^T\right)\\
&= \left(\mathbb{D}\left(\boldsymbol{v}\right),\nabla\boldsymbol{v}\right) + \left(\mathbb{D}\left(\boldsymbol{v}\right),\left(\nabla\boldsymbol{v}\right)^T\right) = 2\left\|\mathbb{D}\left(\boldsymbol{v}\right)\right\|_{L^2(\Omega)}^2.
\end{aligned}
$$

Inserting this identity in (3.44) proves (3.42). ■

Remark 3.38 (On Korn's Inequality) Korn's equality (3.43) provides an estimate of the $L^2(\Omega)$ norm of the deformation tensor of $\boldsymbol{v}$, which is the symmetric part of $\nabla\boldsymbol{v}$, with the $L^2(\Omega)$ norm of the gradient of $\boldsymbol{v}$ if $\boldsymbol{v} \in V$. There are further estimates of this type, e.g., in other Lebesgue spaces and under different conditions on $\boldsymbol{v}$, see Friedrichs (1947), Horgan (1995) for details. It should be also noted that estimates of Korn's type do not hold for all spaces, e.g., they do not hold for certain non-conforming finite element spaces. □

Lemma 3.39 (Boundedness, Continuity, and Norm of the Bilinear Form $b(\cdot,\cdot)$) *The bilinear form $b(\cdot,\cdot)$ from (3.34) is bounded*

$$|b(\boldsymbol{v},q)| \le \|\boldsymbol{v}\|_V \|q\|_Q$$

and consequently it is continuous. In addition, it holds $\|b\| = 1$.

Proof The boundedness follows with the Cauchy–Schwarz inequality (A.10) and (3.41)

$$|b(\boldsymbol{v},q)| = \left|-\int_\Omega (\nabla\cdot\boldsymbol{v})\, q\, d\boldsymbol{x}\right| \le \|\nabla\cdot\boldsymbol{v}\|_{L^2(\Omega)}\|q\|_{L^2(\Omega)} \le \|\nabla\boldsymbol{v}\|_{L^2(\Omega)}\|q\|_{L^2(\Omega)}. \tag{3.45}$$

Continuity follows from boundedness.

The statement concerning the norm of $b(\cdot,\cdot)$ follows from the definition (3.3) of this norm and (3.45). ■

Lemma 3.40 (V_{div} is a Closed Subspace of V) *The subspace of weakly divergence-free functions V_{div} is closed in V.*

Proof The proof is essentially the same as in the general case, see Lemma 3.9. It is given here for completeness of presentation.

Since $b(\cdot,\cdot)$ is a bilinear form, it follows that

$$
\begin{aligned}
&(\alpha\nabla\cdot\boldsymbol{v}_1 + \beta\nabla\cdot\boldsymbol{v}_2, q) = \alpha\left(\nabla\cdot\boldsymbol{v}_1, q\right) + \beta\left(\nabla\cdot\boldsymbol{v}_2, q\right) = 0\\
&\qquad\qquad\qquad\qquad \forall\ \alpha,\beta\in\mathbb{R},\ \boldsymbol{v}_1,\boldsymbol{v}_2\in V_{\text{div}},\ q\in Q.
\end{aligned}
$$

Hence, any linear combination of weakly divergence-free functions is weakly divergence-free and therefore V_{div} is a subspace of V.

Let $\boldsymbol{v} \in V$ be arbitrary such that a sequence $\boldsymbol{v}_n \to \boldsymbol{v}$, $\boldsymbol{v}_n \in V_{\text{div}}$, $n = 1, 2, \ldots$, exists which converges to $\boldsymbol{v}$ in V, i.e., $\|\boldsymbol{v} - \boldsymbol{v}_n\|_V \to 0$ as $n \to \infty$. To prove that V_{div} is closed, one has to show that $\boldsymbol{v} \in V_{\text{div}}$. Let $q \in Q$ be arbitrary but fixed, then it follows from the continuity of $b(\cdot, \cdot)$ that

$$b(\boldsymbol{v}, q) = b\left(\lim_{n\to\infty} \boldsymbol{v}_n, q\right) = \lim_{n\to\infty} b\left(\boldsymbol{v}_n, q\right) = \lim_{n\to\infty} 0 = 0.$$

Since $q \in Q$ was arbitrary, one gets $b(\boldsymbol{v}, q) = 0$ for all $q \in Q$, i.e., $\boldsymbol{v} \in V_{\text{div}}$. ∎

Lemma 3.41 (Isomorphism of the Gradient Operator) *If $\boldsymbol{f} \in V'$ satisfies*

$$\langle \boldsymbol{f}, \boldsymbol{v}\rangle_{V',V} = 0 \quad \forall\ \boldsymbol{v} \in V_{\text{div}},$$

then there exists a unique $q \in Q$ such that

$$\boldsymbol{f} = grad(q).$$

That means, the range of the gradient operator consists of the functionals in V' that vanish on V_{div}

$$\tilde{V}' = \left\{\boldsymbol{f} \in V' \ :\ \langle \boldsymbol{f}, \boldsymbol{v}\rangle_{V',V} = 0, \quad \forall\ \boldsymbol{v} \in V_{\text{div}}\right\},$$

compare (3.11)*, and this operator is an isomorphism from Q onto $\tilde{V}'$.*

Proof It is known that the range of grad is a subspace of V', see Remark 3.29. It can be even shown, see Girault and Raviart (1986, p. 20) on the basis of results from Carroll et al. (1966) or Duvaut and Lions (1972), that range (grad) is a closed subspace of V'. The operators $-$div and grad are dual operators. From the Closed Range Theorem of Banach, Theorem A.71 iv), it follows that range (grad) is the subspace of functionals from V' which vanish on the kernel of div, i.e., range (grad) $= \tilde{V}'$.

To prove uniqueness, consider $q_1, q_2 \in Q$ with

$$\boldsymbol{f} = \text{grad}(q_1) = \text{grad}(q_2).$$

Then, one has

$$\mathbf{0} = \text{grad}(q_1) - \text{grad}(q_2) = \text{grad}(q_1 - q_2).$$

Hence $q_1 - q_2 \in \ker(\text{grad})$, i.e., $q_1 - q_2 \in Q$ is almost everywhere a constant function. The only function that is constant almost everywhere in Q is $q = 0$. It follows that $q_1 = q_2$ in the sense of $L^2(\Omega)$. ∎

Remark 3.42 (Orthogonal Decomposition of V) The space V can be decomposed into orthogonal subspaces

$$V = V_{\text{div}} \oplus V_{\text{div}}^{\perp},$$

where the orthogonality is based on the inner product (3.35) of V. □

Lemma 3.43 (Isomorphism of the Divergence Operator) *The operator div is an isomorphism from* $V_{\mathrm{div}}^{\perp}$ *onto Q.*

Proof The proof is similar to the second part of the proof of Lemma 3.12.

The operator $-\mathrm{div}$ is the dual of grad. From Lemma 3.41 it follows that $-\mathrm{div}$, and with that the operator div, is an isomorphism from the dual space of $\tilde{V}'$ onto Q'. It will be shown that the dual space of $\tilde{V}'$ is $V_{\mathrm{div}}^{\perp}$, which is equivalent to show that $\tilde{V}' = \left(V_{\mathrm{div}}^{\perp}\right)'$. To this end, an isomorphism $\left(V_{\mathrm{div}}^{\perp}\right)' \to \tilde{V}'$ will be constructed.

Let $\tilde{\boldsymbol{g}} \in \left(V_{\mathrm{div}}^{\perp}\right)'$, then a functional $\boldsymbol{g} \in V'$ can be defined by setting

$$\langle \boldsymbol{g}, \boldsymbol{v} \rangle_{V',V} = \left\langle \tilde{\boldsymbol{g}}, \boldsymbol{v}^{\perp} \right\rangle_{V',V} \quad \forall\ \boldsymbol{v} \in V,$$

where $\boldsymbol{v}^{\perp}$ is the orthogonal projection of $\boldsymbol{v}$ onto $V_{\mathrm{div}}^{\perp}$. In particular, it holds for all $\boldsymbol{v} \in V_{\mathrm{div}}$ that

$$\langle \boldsymbol{g}, \boldsymbol{v} \rangle_{V',V} = \langle \tilde{\boldsymbol{g}}, \boldsymbol{0} \rangle_{V',V} = 0.$$

Hence, $\boldsymbol{g} \in \tilde{V}'$. In this way, a linear mapping

$$\left(V_{\mathrm{div}}^{\perp}\right)' \to \tilde{V}', \quad \tilde{\boldsymbol{g}} \mapsto \boldsymbol{g}$$

is defined.

First, it will be shown that this mapping is injective. Let $\tilde{\boldsymbol{g}}_1, \tilde{\boldsymbol{g}}_2 \in \left(V_{\mathrm{div}}^{\perp}\right)'$ with

$$\langle \boldsymbol{g}, \boldsymbol{v} \rangle_{V',V} = \left\langle \tilde{\boldsymbol{g}}_1, \boldsymbol{v}^{\perp} \right\rangle_{V',V} = \left\langle \tilde{\boldsymbol{g}}_2, \boldsymbol{v}^{\perp} \right\rangle_{V',V} \quad \forall\ \boldsymbol{v} \in V,$$

then it is

$$\left\langle \tilde{\boldsymbol{g}}_1 - \tilde{\boldsymbol{g}}_2, \boldsymbol{v}^{\perp} \right\rangle_{V',V} = 0 \quad \forall\ \boldsymbol{v} \in V.$$

This equality holds in particular for all $\boldsymbol{v} \in V_{\mathrm{div}}^{\perp}$, from which it follows that the functionals $\tilde{\boldsymbol{g}}_1, \tilde{\boldsymbol{g}}_2$ are identical.

Next, the surjectivity of the mapping will be proved. Let $\boldsymbol{g} \in \tilde{V}'$, i.e., $\langle \boldsymbol{g}, \boldsymbol{v} \rangle_{V',V} = 0$ for all $\boldsymbol{v} \in V_{\mathrm{div}}$. Consider an arbitrary $\boldsymbol{v} \in V$. This function can be decomposed into $\boldsymbol{v} = \boldsymbol{v}_{\mathrm{div}} + \boldsymbol{v}_{\mathrm{div}}^{\perp}$ with $\boldsymbol{v}_{\mathrm{div}} \in V_{\mathrm{div}}$, $\boldsymbol{v}_{\mathrm{div}}^{\perp} \in V_{\mathrm{div}}^{\perp}$. Since $\boldsymbol{v}$ is arbitrary, also $\boldsymbol{v}_{\mathrm{div}}^{\perp}$ is arbitrary. It follows that

$$\langle \boldsymbol{g}, \boldsymbol{v} \rangle_{V',V} = \langle \boldsymbol{g}, \boldsymbol{v}_{\mathrm{div}} \rangle_{V',V} + \left\langle \boldsymbol{g}, \boldsymbol{v}_{\mathrm{div}}^{\perp} \right\rangle_{V',V} = \left\langle \boldsymbol{g}, \boldsymbol{v}_{\mathrm{div}}^{\perp} \right\rangle_{V',V} \quad \forall\ \boldsymbol{v}_{\mathrm{div}}^{\perp} \in V_{\mathrm{div}}^{\perp}.$$

This relation defines a functional on $V_{\mathrm{div}}^{\perp}$ which is mapped onto $\boldsymbol{g}$. Consequently, the mapping is surjective. ■

Corollary 3.44 (Each Pressure is the Divergence of a Velocity Field) *For each $q \in Q$ there is a unique $\boldsymbol{v} \in V_{\text{div}}^{\perp} \subset V$ such that*

$$\nabla \cdot \boldsymbol{v} = q \quad \textit{and} \quad \|q\|_Q \le \|\boldsymbol{v}\|_V\,, \quad \|\boldsymbol{v}\|_V \le C\,\|q\|_Q\,, \tag{3.46}$$

with C independent of $\boldsymbol{v}$ and q. In the proof of Theorem 3.46 below, it will be shown that $C = \beta_{\text{is}}^{-1}$.

Proof The existence of a unique $\boldsymbol{v} \in V_{\text{div}}^{\perp}$ with $\nabla \cdot \boldsymbol{v} = q$ follows from the isomorphism of the divergence operator, see Lemma 3.43. Then, one gets with (3.41)

$$\|q\|_Q = \|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)} \le \|\nabla \boldsymbol{v}\|_{L^2(\Omega)} = \|\boldsymbol{v}\|_V\,.$$

The inverse map of the divergence operator is an isomorphism, too. In particular, it is bounded, see Theorem A.70. Hence there is a $C > 0$ such that $\|\boldsymbol{v}\|_V = \left\|\text{div}^{-1} q\right\|_V \le C\,\|q\|_Q$ for all $q \in Q$ and all $\boldsymbol{v} \in V_{\text{div}}^{\perp}$. ■

Remark 3.45 (Forms of the Inf-Sup Condition (3.14) *Found in the Literature)* Since with each function which can be inserted in the inf-sup condition also its negative can be inserted, one has

$$\begin{aligned}
&\inf_{q\in Q, q\neq 0}\ \sup_{\boldsymbol{v}\in V, \boldsymbol{v}\neq \boldsymbol{0}} \frac{b(\boldsymbol{v}, q)}{\|\boldsymbol{v}\|_V\,\|q\|_Q} \\
&= \inf_{q\in Q, q\neq 0}\ \sup_{\boldsymbol{v}\in V, \boldsymbol{v}\neq \boldsymbol{0}} \frac{-\left(\nabla \cdot \boldsymbol{v}, q\right)}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)}\,\|q\|_{L^2(\Omega)}} \\
&= \inf_{q\in Q, q\neq 0}\ \sup_{\boldsymbol{v}\in V, \boldsymbol{v}\neq \boldsymbol{0}} \frac{\left(\nabla \cdot \boldsymbol{v}, q\right)}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)}\,\|q\|_{L^2(\Omega)}} \ge \beta_{\text{is}} > 0.
\end{aligned}$$

The last line is a form that can be found often in the literature.

Another form is that for each $q \in Q$, it holds that

$$\sup_{\boldsymbol{v}\in V, \boldsymbol{v}\neq \boldsymbol{0}} \frac{\left(\nabla \cdot \boldsymbol{v}, q\right)}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)}} \ge \beta_{\text{is}}\,\|q\|_{L^2(\Omega)}\,. \tag{3.47}$$

□

Theorem 3.46 (Inf-Sup Condition for V and Q) *The spaces V and Q satisfy the inf-sup condition* (3.14)*, i.e., there is a $\beta_{\text{is}} > 0$ such that*

$$\inf_{q\in Q, q\neq 0}\ \sup_{\boldsymbol{v}\in V, \boldsymbol{v}\neq \boldsymbol{0}} \frac{\left(\nabla \cdot \boldsymbol{v}, q\right)}{\|\boldsymbol{v}\|_V\,\|q\|_Q} \ge \beta_{\text{is}}. \tag{3.48}$$

Proof Let $q \in Q$ be arbitrary. By Corollary 3.44 there exists a unique $\boldsymbol{v} \in V_{\mathrm{div}}^{\perp}$ such that

$$\nabla \cdot \boldsymbol{v} = q, \quad \|\boldsymbol{v}\|_V \le C \|q\|_Q .$$

It follows that

$$\frac{(\nabla \cdot \boldsymbol{v}, q)}{\|\boldsymbol{v}\|_V} = \frac{(q, q)}{\|\boldsymbol{v}\|_V} = \frac{\|q\|_Q^2}{\|\boldsymbol{v}\|_V} \ge \frac{1}{C} \|q\|_Q .$$

Hence

$$\sup_{\boldsymbol{v} \in V, \boldsymbol{v} \neq \boldsymbol{0}} \frac{(\nabla \cdot \boldsymbol{v}, q)}{\|\boldsymbol{v}\|_V} \ge \frac{1}{C} \|q\|_Q ,$$

and because $q \in Q$ is arbitrary, one obtains

$$\inf_{q \in Q, q \neq 0} \sup_{\boldsymbol{v} \in V, \boldsymbol{v} \neq \boldsymbol{0}} \frac{(\nabla \cdot \boldsymbol{v}, q)}{\|\boldsymbol{v}\|_V \|q\|_Q} \ge \frac{1}{C} =: \beta_{\mathrm{is}}.$$

■

Corollary 3.47 (Estimating the Gradient by the Divergence for Functions from $V_{\mathrm{div}}^{\perp}$) *For all $\boldsymbol{v} \in V_{\mathrm{div}}^{\perp}$ it holds*

$$\|\boldsymbol{v}\|_V \le \frac{1}{\beta_{\mathrm{is}}} \|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)} , \tag{3.49}$$

cf. Lemma 3.12 and (3.16)*.*

Proof From (3.46) and the specification of C, it follows that

$$\|\boldsymbol{v}\|_V \le C \|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)} = \frac{1}{\beta_{\mathrm{is}}} \|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)} .$$

■

Lemma 3.48 (Upper Bound for the Inf-Sup Constant) *It is $\beta_{\mathrm{is}} \le 1$.*

Proof Using Corollary 3.44, one can take $q = \nabla \cdot \boldsymbol{v}$ in the inf-sup condition (3.48). Applying then estimate (3.41) yields

$$\begin{aligned} \beta_{\mathrm{is}} &\le \sup_{\boldsymbol{v} \in V, \boldsymbol{v} \neq \boldsymbol{0}} \frac{(\nabla \cdot \boldsymbol{v}, \nabla \cdot \boldsymbol{v})}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)}} = \sup_{\boldsymbol{v} \in V, \boldsymbol{v} \neq \boldsymbol{0}} \frac{\|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)}^2}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)}} \\ &\le \sup_{\boldsymbol{v} \in V, \boldsymbol{v} \neq \boldsymbol{0}} \frac{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)}}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)}} = 1. \end{aligned}$$

■

Remark 3.49 (Dependency of β_{is} on the Domain)

- Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be a bounded domain that is star-shaped with respect to some ball with radius $r > 0$. Then it holds

$$\frac{1}{\beta_{\mathrm{is}}} \le C \left(\frac{\mathrm{diam}(\Omega)}{r}\right)^d \left(1 + \frac{\mathrm{diam}(\Omega)}{r}\right), \tag{3.50}$$

  where C is a constant depending only on the dimension. Inequality (3.50) is a special case of the estimate proved in Galdi (2011, Theorem III.3.1).
- The dependency of β_{is} on Ω is analyzed for two types of two-dimensional domains in Chizhonkov and Olshanskii (2000). In particular, for a rectangular domain with side lengths $l_1 \le l_2$, it is proved that $\beta_{\mathrm{is}} = \mathcal{O}(l_1/l_2)$. Consequently, the inf-sup constant is very small in long and thin channel-type domains.

□

3.3 General Considerations on Appropriate Function Spaces for Finite Element Discretizations

Remark 3.50 (On Finite Element Methods) A brief introduction to finite element methods is provided in Appendix B. The main idea of using finite element methods consists in replacing the infinite-dimensional spaces V and Q by a finite-dimensional velocity space V^h and a finite-dimensional pressure space Q^h and to apply the Galerkin method, see Remark B.10. If $V^h \subset V$ and $Q^h \subset Q$, the finite element method is called conforming, otherwise it is called non-conforming.

For incompressible flow problems, the pair of velocity-pressure finite element spaces is denoted by V^h/Q^h. It is usual that it will not be emphasized in the notation that V^h consists of vector-valued functions and that Q^h is possibly intersected with $L_0^2(\Omega)$, depending on the boundary condition. □

Remark 3.51 (Application of the Abstract Theory, the Discrete Inf-Sup Condition) Clearly, the finite-dimensional spaces are Hilbert spaces and the theory developed in Sect. 3.1 can be applied for the investigation of the existence and the uniqueness of a solution of the finite element problems arising in the discretization of incompressible flow models. In particular, the spaces V^h and Q^h have to satisfy an inf-sup condition of the form

$$\inf_{q^h \in Q^h \setminus \{0\}} \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b^h\left(\boldsymbol{v}^h, q^h\right)}{\|\boldsymbol{v}^h\|_{V^h} \|q^h\|_{Q^h}} \ge \beta_{\mathrm{is}}^h > 0 \tag{3.51}$$

or equivalently that there is a $\beta_{\mathrm{is}}^h > 0$ such that

$$\sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b^h\left(\boldsymbol{v}^h, q^h\right)}{\|\boldsymbol{v}^h\|_{V^h}} \ge \beta_{\mathrm{is}}^h \|q^h\|_{Q^h} \quad \forall\, q^h \in Q^h. \tag{3.52}$$

This condition is called discrete inf-sup condition. In (3.51) and (3.52), the bilinear form $b^h : V^h \times Q^h \to \mathbb{R}$ is defined by

$$b^h\left(\boldsymbol{v}^h, q^h\right) = -\sum_{K\in\mathcal{T}^h} \left(\nabla\cdot\boldsymbol{v}^h, q^h\right)_K, \tag{3.53}$$

where $\mathcal{T}^h$ is a triangulation of Ω and $K \in \mathcal{T}^h$ are the mesh cells. For conforming finite element spaces, the bilinear form $b^h(\cdot,\cdot)$ can be written in the same form as the bilinear form $b(\cdot,\cdot)$ with an integral on Ω, see (3.34). In this case, $b^h(\cdot,\cdot)$ is just the restriction of $b(\cdot,\cdot)$ from $V \times Q$ to $V^h \times Q^h$. The norms in the denominator are defined by

$$\left\|\boldsymbol{v}^h\right\|_{V^h} = \left(\sum_{K\in\mathcal{T}^h} \left(\nabla\boldsymbol{v}^h, \nabla\boldsymbol{v}^h\right)_K\right)^{1/2}, \quad \left\|q^h\right\|_{Q^h} = \left\|q^h\right\|_{L^2(\Omega)}. \tag{3.54}$$

For a conforming velocity finite element space, it is $\left\|\boldsymbol{v}^h\right\|_{V^h} = \left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)}$.

In the same way as in the proof of Lemma 3.48, one finds for conforming finite element spaces that $\beta_{\mathrm{is}}^h \le 1$. □

Remark 3.52 (Non-inheritance of the Inf-Sup Condition from V and Q) Consider a conforming finite element method, then

$$\sup_{\boldsymbol{v}^h\in V^h\setminus\{\boldsymbol{0}\}} \frac{b\left(\boldsymbol{v}^h, q\right)}{\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)} \left\|q\right\|_{L^2(\Omega)}} \le \sup_{\boldsymbol{v}\in V\setminus\{\boldsymbol{0}\}} \frac{b\left(\boldsymbol{v}, q\right)}{\left\|\nabla\boldsymbol{v}\right\|_{L^2(\Omega)} \left\|q\right\|_{L^2(\Omega)}}$$

since the supremum in V^h is searched in a smaller set. In general, the strong inequality will hold. Hence

$$\begin{aligned} &\inf_{q\in Q\setminus\{0\}} \sup_{\boldsymbol{v}^h\in V^h\setminus\{\boldsymbol{0}\}} \frac{b\left(\boldsymbol{v}^h, q\right)}{\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)} \left\|q\right\|_{L^2(\Omega)}} \\ &\le \inf_{q\in Q\setminus\{0\}} \sup_{\boldsymbol{v}\in V\setminus\{\boldsymbol{0}\}} \frac{b\left(\boldsymbol{v}, q\right)}{\left\|\nabla\boldsymbol{v}\right\|_{L^2(\Omega)} \left\|q\right\|_{L^2(\Omega)}} \end{aligned} \tag{3.55}$$

and the continuous inf-sup parameter β_{is}, which is a lower bound of the right-hand side of (3.55), cannot be expected to be a lower bound of the left-hand side, too. In fact, the left-hand side is zero since V^h and Q do not satisfy an inf-sup condition, see Remark 3.7. In this remark, it was discussed that the dimension of the pressure space should not exceed the dimension of the velocity space in order to get a well-posed problem.

Turning to a finite element method, the infinite-dimensional space Q has to be replaced by a finite-dimensional space Q^h. This replacement might lead to an increase of the left-hand side of (3.55) since now the infimum is taken in a smaller set. Eventually, Q^h becomes sufficiently small such that

$$\inf_{q^h\in Q^h\setminus\{0\}} \sup_{\boldsymbol{v}^h\in V^h\setminus\{\boldsymbol{0}\}} \frac{b\left(\boldsymbol{v}^h, q^h\right)}{\|\nabla \boldsymbol{v}^h\|_{L^2(\Omega)} \|q^h\|_{L^2(\Omega)}}$$

becomes positive. Then, V^h and Q^h satisfy a discrete inf-sup condition.

These considerations give a rough idea about appropriate choices for the finite element spaces with respect to the discrete inf-sup condition. The velocity space V^h should be sufficiently large such that the supremum of $\boldsymbol{v}^h \in V^h$ becomes large and the pressure space Q^h should be sufficiently small such that the infimum of $q^h \in Q^h$ becomes large, too. A condition in this direction can be found already in Remark 3.7, where $n_Q \le n_V$ was required. However, there is a conflicting requirement for the pressure finite element space. For obtaining accurate results, this space has to be large enough such that it is possible to approximate the continuous pressure sufficiently well. Also an accurate conservation of mass requires a large discrete pressure space compared with the discrete velocity space, see Remark 3.56 for details. □

Lemma 3.53 ($\beta_{\text{is}}^h \le \beta_{\text{is}}$ for Conforming Finite Element Spaces) *Consider a family of finite element spaces $\left\{V^h \times Q^h\right\}$ with $V^h \subset V$, $Q^h \subset Q$, and let this family satisfy the discrete inf-sup condition (3.51) independently of h. Assume that for each $q \in Q \cap H^1(\Omega)$ there is a $q^h \in Q^h$ such that*

$$\left\|q - q^h\right\|_{L^2(\Omega)} \le Ch \|q\|_{H^1(\Omega)}, \tag{3.56}$$

with C independent of q and h. Then, it holds $\beta_{\text{is}}^h \le \beta_{\text{is}}$, where both values are the largest possible values in (3.51) and (3.48), respectively.

Proof The proof follows Chizhonkov and Olshanskii (2000). Since $q \in Q \cap H^1(\Omega) \subset Q$, it is

$$\inf_{q\in\left(Q\cap H^1(\Omega)\right)\setminus\{0\}} \sup_{\boldsymbol{v}\in V\setminus\{\boldsymbol{0}\}} \frac{(\nabla\cdot\boldsymbol{v}, q)}{\|\boldsymbol{v}\|_V \|q\|_Q} \ge \inf_{q\in Q\setminus\{0\}} \sup_{\boldsymbol{v}\in V\setminus\{\boldsymbol{0}\}} \frac{(\nabla\cdot\boldsymbol{v}, q)}{\|\boldsymbol{v}\|_V \|q\|_Q}.$$

On the other hand, because of the density of $Q \cap H^1(\Omega)$ in Q, which follows from Theorem A.38, even the equal sign holds, such that

$$\beta_{\text{is}} = \inf_{q\in\left(Q\cap H^1(\Omega)\right)\setminus\{0\}} \sup_{\boldsymbol{v}\in V\setminus\{\boldsymbol{0}\}} \frac{(\nabla\cdot\boldsymbol{v}, q)}{\|\boldsymbol{v}\|_V \|q\|_Q}. \tag{3.57}$$

Consider an arbitrary $q \in Q \cap H^1(\Omega)$ and $\varepsilon \in (0, 1)$. For sufficiently small h, one has

$$Ch \|q\|_{H^1(\Omega)} \le \varepsilon \|q\|_{L^2(\Omega)},$$

such that (3.56) gives

$$\left\|q - q^h\right\|_{L^2(\Omega)} \leq \varepsilon \|q\|_{L^2(\Omega)} . \tag{3.58}$$

By the triangle inequality, one obtains from this relation

$$\|q\|_{L^2(\Omega)} \leq \left\|q - q^h\right\|_{L^2(\Omega)} + \left\|q^h\right\|_{L^2(\Omega)} \leq \varepsilon \|q\|_{L^2(\Omega)} + \left\|q^h\right\|_{L^2(\Omega)} ,$$

which is equivalent to

$$(1 - \varepsilon) \|q\|_{L^2(\Omega)} \leq \left\|q^h\right\|_{L^2(\Omega)} . \tag{3.59}$$

For each $q^h \in Q^h$, one gets with the discrete inf-sup condition (3.51), the property that the supremum of a sum is lower or equal than the sum of the suprema, the Cauchy–Schwarz inequality (A.10), estimates (3.41), (3.58), (3.59) with $q \in Q \cap H^1(\Omega)$, and the inclusion $V^h \subset V$

$$\begin{aligned}
\beta_{\text{is}}^h &\leq \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b\left(\boldsymbol{v}^h, q^h\right)}{\|\nabla \boldsymbol{v}^h\|_{L^2(\Omega)} \|q^h\|_{L^2(\Omega)}} \\
&\leq \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b\left(\boldsymbol{v}^h, q\right)}{\|\nabla \boldsymbol{v}^h\|_{L^2(\Omega)} \|q^h\|_{L^2(\Omega)}} + \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b\left(\boldsymbol{v}^h, q^h - q\right)}{\|\nabla \boldsymbol{v}^h\|_{L^2(\Omega)} \|q^h\|_{L^2(\Omega)}} \\
&\leq \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b\left(\boldsymbol{v}^h, q\right)}{\|\nabla \boldsymbol{v}^h\|_{L^2(\Omega)} \|q^h\|_{L^2(\Omega)}} + \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{\left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)} \left\|q - q^h\right\|_{L^2(\Omega)}}{\|\nabla \boldsymbol{v}^h\|_{L^2(\Omega)} \|q^h\|_{L^2(\Omega)}} \\
&\leq \frac{1}{1 - \varepsilon} \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b\left(\boldsymbol{v}^h, q\right)}{\|\nabla \boldsymbol{v}^h\|_{L^2(\Omega)} \|q\|_{L^2(\Omega)}} + \frac{\varepsilon}{1 - \varepsilon} \\
&\leq \frac{1}{1 - \varepsilon} \sup_{\boldsymbol{v} \in V \setminus \{\boldsymbol{0}\}} \frac{b\left(\boldsymbol{v}, q\right)}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \|q\|_{L^2(\Omega)}} + \frac{\varepsilon}{1 - \varepsilon}.
\end{aligned}$$

Taking the infimum with respect to q on both sides of this inequality gives with (3.57)

$$\beta_{\text{is}}^h \leq \frac{1}{1 - \varepsilon} \left(\beta_{\text{is}} + \varepsilon\right) \quad \Longleftrightarrow \quad \beta_{\text{is}}^h - \beta_{\text{is}} \leq \varepsilon \left(1 + \beta_{\text{is}}^h\right).$$

Since the right-hand side of the last inequality is arbitrarily close to zero for sufficiently small ε, the relation $\beta_{\text{is}}^h > \beta_{\text{is}}$ cannot hold, which proves the statement of the lemma. ■

Remark 3.54 (The Babuška Condition) The discrete inf-sup condition (3.51) is also called discrete Babuška–Brezzi or discrete Ladyzhenskaya–Babuška–Brezzi (LBB) condition.

To be more precise, Babuška (1971) considered the discrete inf-sup condition

$$\inf_{\substack{(\boldsymbol{u}^h,p^h)\in V^h\times Q^h\\(\boldsymbol{u}^h,p^h)\neq(\boldsymbol{0},0)}}\sup_{\substack{(\boldsymbol{v}^h,q^h)\in V^h\times Q^h\\(\boldsymbol{v}^h,q^h)\neq(\boldsymbol{0},0)}}\frac{a\left(\boldsymbol{u}^h,\boldsymbol{v}^h\right)+b\left(\boldsymbol{v}^h,p^h\right)+b\left(\boldsymbol{u}^h,q^h\right)}{\left(\|\boldsymbol{u}^h\|_{V^h}+\|p^h\|_{Q^h}\right)\left(\|\boldsymbol{v}^h\|_{V^h}+\|q^h\|_{Q^h}\right)}\geq\tilde{\beta}_{\mathrm{is}}^h>0. \qquad (3.60)$$

It can be shown that the discrete inf-sup condition (3.51), together with appropriate assumptions on the bilinear form $a(\cdot,\cdot)$, and the discrete Babuška condition imply each other, see Xu and Zikatanov (2003), Demkowicz (2006), where the inf-sup constants are in general different. In Braess and Blömer (1990), it is proved that an inf-sup condition of form (3.60) for a generalized saddle point problem follows from the satisfaction of an inf-sup condition of type (3.51), even in the continuous setting.

From the fulfillment of the discrete Babuška inf-sup condition (3.60), the unique solvability of the coupled discrete system can be concluded. This approach can be used if the bilinear form $a(\cdot,\cdot)$ becomes complicated, e.g., if so-called stabilized methods are used, see Corollary 5.40. □

Remark 3.55 (The Space of Discretely Divergence-free Functions) Exactly as in Sect. 3.1, a linear operator B^h can be associated with the bilinear form $b^h(\cdot,\cdot)$

$$B^h\ :\ V^h\to\left(Q^h\right)',\quad \left\langle B^h\boldsymbol{v}^h,q^h\right\rangle_{(Q^h)',Q^h}=b^h\left(\boldsymbol{v}^h,q^h\right). \qquad (3.61)$$

Thus, B^h is a discrete (negative) divergence operator div^h. Note that by the representation theorem of Riesz, Theorem B.3, the space $\left(Q^h\right)'$ can be identified with Q^h. Usually, it is $\nabla\cdot\boldsymbol{v}^h\notin Q^h$. Thus, definition (3.61) strictly speaking uses the $L^2(\Omega)$ projection of $\nabla\cdot\boldsymbol{v}^h$ into Q^h, which reads for a conforming finite element method

$$\left(B^h\boldsymbol{v}^h,q^h\right)=-\left(P_{L^2}^h\left(\nabla\cdot\boldsymbol{v}^h\right),q^h\right)=-\left(\nabla\cdot\boldsymbol{v}^h,q^h\right)\quad\forall\ q^h\in Q^h.$$

From Sect. 3.1 it is known that the kernel of B^h plays an important role in the theory. This kernel is called the space of discretely divergence-free functions

$$V_{\mathrm{div}}^h=\left\{\boldsymbol{v}^h\in V^h\ :\ b^h\left(\boldsymbol{v}^h,q^h\right)=0\ \forall\ q^h\in Q^h\right\}. \qquad (3.62)$$

The dual operator of the discrete divergence is a discrete gradient operator

$$\left(B^h\right)^T\ :\ Q^h\to\left(V^h\right)'\quad\left\langle\left(B^h\right)^Tq^h,\boldsymbol{v}^h\right\rangle_{(V^h)',V^h}=b^h\left(\boldsymbol{v}^h,q^h\right), \qquad (3.63)$$

which will be denoted by grad^h.

Since the discrete divergence B^h is a linear operator between finite-dimensional spaces, it can be represented by a matrix, once bases in V^h and Q^h have been chosen. This matrix has the dimension $\dim Q^h \times \dim V^h$. The notation in (3.63) for the discrete gradient is used because it can be represented with the transposed matrix. By the Riesz representation theorem, Theorem B.3, $\left(V^h\right)'$ can be identified with V^h. In particular it holds that $\dim\left(V^h\right) = \dim\left(\left(V^h\right)'\right)$. □

Remark 3.56 (On Discretely Divergence-free Functions, Violation of Mass Conservation) Let $Q^h \subsetneq Q$, then the functions from V^h_{div} need to satisfy less conditions than the functions from V_{div}. Consequently, there is no injection, i.e., in general $V^h_{\mathrm{div}} \not\subset V_{\mathrm{div}}$. In particular, one finds that discretely divergence-free functions are in general neither weakly nor pointwise divergence-free. Thus, the conservation of mass, which was modeled by the divergence-free constraint, Sect. 2.1, is not satisfied exactly, but only in some approximate or mean sense.

When applying finite element methods for the simulation of incompressible flows, one has to be aware that the conservation of mass might be violated. The extent of the violation depends on the concrete choice of the finite element spaces. Note that there are some pairs of finite element spaces which are mass conservative, like the Scott–Vogelius finite element, see Sect. 3.6.3. The topic of mass conservation will be discussed in detail in Sect. 4.6.

Consider the case $V^h \subset V$. Note that a finite element function $\boldsymbol{v}^h \in V^h_{\mathrm{div}}$ is weakly divergence-free if $\nabla \cdot V^h \subseteq Q^h$. In this case, it is $\nabla \cdot \boldsymbol{v}^h \in Q^h$ such that from the definition (3.62) of V^h_{div} it follows that

$$0 = b\left(\boldsymbol{v}^h, \nabla \cdot \boldsymbol{v}^h\right) = \left\|\nabla \cdot \boldsymbol{v}^h\right\|^2_{L^2(\Omega)}.$$

Thus, the divergence vanishes in the sense of $L^2(\Omega)$. For the condition $\nabla \cdot V^h \subseteq Q^h$ to be hold, Q^h has to be sufficiently large or V^h should be sufficiently small. These requirements are just contrary to the requirements for the fulfillment of the discrete inf-sup condition, see the discussion at the end of Remark 3.52. Thus, one might suspect that the enforcement of the discrete inf-sup condition (3.51) probably has to be paid with a relaxation of the continuity constraint, as it is in fact the case for most inf-sup stable pairs of finite element spaces. □

Remark 3.57 (The Discrete Inf-Sup Parameter β^h_{is}) A standard approach of discretizing partial differential equations consists in starting with a coarse triangulation of Ω, solving the considered problem on this triangulation, refining the grid, and repeating this process until, e.g., a finest grid is reached on which the solution is sufficiently accurate, or on which memory restrictions prevent a further refinement. On all grid levels, finite element spaces which satisfy the discrete inf-sup condition (3.51) should be used, where the corresponding inf-sup parameters β^h_{is} might be different.

Finite element error analysis will reveal that the inf-sup parameters enter the error estimates, e.g., see Theorem 4.21 for the Stokes equations. The error bounds depend on inverse of powers of β_{is}^h. Thus, a behavior of the form $\beta_{\text{is}}^h \to 0$ for successive refinements leads to a deterioration of the order of convergence in the error estimates, e.g., compare Remark 4.29. For this reason, it is important that the used finite element spaces satisfy (3.51) with a parameter $\beta_{\text{is}}^h > 0$ that is independent of the refinement level of the grid or, equivalently, independent of the mesh size parameter h. □

Lemma 3.58 (Each Discrete Pressure is the Divergence of a Discrete Velocity Field) *Let $V^h \subset V$ with $V_h = V_{\text{div}}^h \oplus \left(V_{\text{div}}^h\right)^\perp$ and let the discrete inf-sup condition* (3.51) *be satisfied. Then there is for each $q^h \in Q^h$ a unique $\boldsymbol{v}^h \in \left(V_{\text{div}}^h\right)^\perp$ such that*

$$\nabla \cdot \boldsymbol{v}^h = q^h, \quad \left\|\boldsymbol{v}^h\right\|_V \le \frac{1}{\beta_{\text{is}}^h} \left\|q^h\right\|_Q. \tag{3.64}$$

Proof By Lemma 3.12 it is known that the satisfaction of the discrete inf-sup condition, point i) in Lemma 3.12, is equivalent with the existence of an isomorphism between $\left(V_{\text{div}}^h\right)^\perp$ and Q^h and with the inequality from (3.64), point iii) of Lemma 3.12. ■

Remark 3.59 (Importance of the Best Approximation Error) In the Galerkin method, the error of the finite element solution $u^h \in V^h$ to the solution of the continuous problem $u \in V$, in the norm of V, can be estimated with the best approximation error, see the Lemma of Cea, Lemma B.12. For incompressible flow problems, it is often convenient to perform the finite element error analysis in the space V_{div}^h, since in V_{div}^h the problem is only an equation for the velocity and not a coupled system. Sometimes, it turns out that the best approximation error in V_{div}^h can be estimated directly, e.g., by constructing a sequence of elements in V_{div}^h which have the optimal order of convergence. One example where this can be done is the non-conforming Crouzeix–Raviart element P_1^{nc}/P_0, see Lemma 4.53. However, estimates of the best approximation error are generally known only for standard finite element spaces, which can be used for V^h, e.g., see the interpolation error estimates in Appendix C. With the help of the discrete inf-sup condition, it is possible to estimate the best approximation error in V_{div}^h with the best approximation error in V^h. □

Lemma 3.60 (Best Approximation Estimate for V_{div}^h) *Let $V^h \subset V$, $\boldsymbol{v} \in V_{\text{div}}$, and let the discrete inf-sup condition* (3.51) *hold. Then*

$$\inf_{\boldsymbol{v}^h \in V_{\text{div}}^h} \left\|\nabla \left(\boldsymbol{v} - \boldsymbol{v}^h\right)\right\|_{L^2(\Omega)} \le \left(1 + \frac{1}{\beta_{\text{is}}^h}\right) \inf_{\boldsymbol{w}^h \in V^h} \left\|\nabla \left(\boldsymbol{v} - \boldsymbol{w}^h\right)\right\|_{L^2(\Omega)}. \tag{3.65}$$

Proof Let $\boldsymbol{w}^h \in V^h$ be arbitrary. Since the discrete inf-sup condition holds, V_{div}^h is not empty, see Remark 3.16. It follows from Hilbert space theory that there is a unique decomposition of $\boldsymbol{w}^h = \boldsymbol{v}^h - \boldsymbol{z}^h$ into a component $\boldsymbol{v}^h \in V_{\text{div}}^h$ and a component

$-z^h \in \left(V_{\text{div}}^h\right)^{\perp}$. Hence, one gets, with $b\left(v^h, q^h\right) = 0$,

$$b\left(z^h, q^h\right) = b\left(v^h - w^h, q^h\right) = b\left(v - w^h, q^h\right) \quad \forall\ q^h \in Q^h. \tag{3.66}$$

Note that $b\left(v, q^h\right) = 0$ since v is weakly divergence-free. From Lemma 3.58, it follows that there is a $q^h = \nabla \cdot z^h \in Q^h$. Inserting this function in (3.66) gives, together with the Cauchy–Schwarz inequality (A.10) and (3.41),

$$\begin{aligned}\left\|\nabla \cdot z^h\right\|_{L^2(\Omega)}^2 &\le \left\|\nabla \cdot \left(v - w^h\right)\right\|_{L^2(\Omega)} \left\|\nabla \cdot z^h\right\|_{L^2(\Omega)} \\ &\le \left\|\nabla \left(v - w^h\right)\right\|_{L^2(\Omega)} \left\|\nabla \cdot z^h\right\|_{L^2(\Omega)}.\end{aligned}$$

With (3.64), one obtains

$$\left\|\nabla z^h\right\|_{L^2(\Omega)} \le \frac{1}{\beta_{\text{is}}^h} \left\|q^h\right\|_{L^2(\Omega)} = \frac{1}{\beta_{\text{is}}^h} \left\|\nabla \cdot z^h\right\|_{L^2(\Omega)} \le \frac{1}{\beta_{\text{is}}^h} \left\|\nabla \left(v - w^h\right)\right\|_{L^2(\Omega)}.$$

Applying the triangle inequality and inserting this estimate yields

$$\begin{aligned}\left\|\nabla \left(v - v^h\right)\right\|_{L^2(\Omega)} &\le \left\|\nabla \left(v - w^h\right)\right\|_{L^2(\Omega)} + \left\|\nabla z^h\right\|_{L^2(\Omega)} \\ &\le \left(1 + \frac{1}{\beta_{\text{is}}^h}\right) \left\|\nabla \left(v - w^h\right)\right\|_{L^2(\Omega)}.\end{aligned}$$

Since w^h was chosen to be arbitrary, one can find for each $w^h \in V^h$ a function $v^h \in V_{\text{div}}^h$ such that this estimate holds, which finishes the proof of the lemma. ■

Remark 3.61 (On Estimate (3.65)) Estimate (3.65) is a worst case estimate. Taking $v^h = \mathbf{0} \in V_{\text{div}}^h$ shows that the left-hand side is always bounded, even if $\beta_{\text{is}}^h = 0$. In contrast, the right-hand side becomes unbounded for $\beta_{\text{is}}^h = 0$ or if β_{is}^h converges sufficiently fast to 0 as $h \to 0$. □

Remark 3.62 (Alternative Estimate for the Best Approximation Error for V_{div}^h in the Case $v \in V_{\text{div}}$) In Girault and Scott (2003), an interpolation operator was constructed that offers an alternative estimate for the best approximation error for V_{div}^h. This interpolation operator is a quasi-local operator that preserves the discrete divergence.

Assume that there is an interpolation operator $I^h \in \mathcal{L}\left(V, V^h\right)$ with

$$\int_K \nabla \cdot \left(v - I^h v\right)\, dx = 0 \qquad \forall\ v \in V,\ K \in \mathcal{T}^h, \tag{3.67}$$

and, compare (C.7),

$$\left\|D^k(v - I^h v)\right\|_{L^p(K)} \le C h_K^{m+1-k} \left\|D^{m+1} v\right\|_{L^p(\omega(K))}, \tag{3.68}$$

for $k \in \{0, 1\}$, $p \in [1, \infty]$, $0 \leq m \leq l$, where l is the polynomial degree of V^h, and C is independent of h and $\omega(K)$. Here, $\omega(K)$ is a neighborhood of K containing all mesh cells that share a face, an edge, or a vertex with K. Then, an operator $I_{\mathrm{div}}^h \in \mathcal{L}\left(V, V^h\right)$ was constructed in Girault and Scott (2003), adding an appropriate correction to I^h such that

$$\left(\nabla \cdot \left(\boldsymbol{v} - I_{\mathrm{div}}^h \boldsymbol{v}\right), q^h\right) = 0 \quad \forall\, q^h \in Q^h, \tag{3.69}$$

and

$$\left\| D^k(\boldsymbol{v} - I_{\mathrm{div}}^h \boldsymbol{v}) \right\|_{L^p(K)} \leq C h_K^{m+1-k} \left\| D^{m+1} \boldsymbol{v} \right\|_{L^p(M_K)}, \tag{3.70}$$

where M_K is a suitable macroelements containing K and C is independent of h, K, and the number of macroelements. The parameters in (3.70) can be chosen in the same way as in (3.68).

The macroelements are allowed to overlap. It is assumed that there is a maximal number, independent of h, of mesh cells contained in one macroelement and that the number of macroelements where a mesh cell belongs to should be bounded, also independently of h.

If $\boldsymbol{v} \in V_{\mathrm{div}}$, it follows that $I_{\mathrm{div}}^h \boldsymbol{v} \in V_{\mathrm{div}}^h$ and with (3.70) for $p = 2$, $k = 1$, one obtains

$$\inf_{\boldsymbol{v}^h \in V_{\mathrm{div}}^h} \left\| \nabla \left(\boldsymbol{v} - \boldsymbol{v}^h\right) \right\|_{L^2(\Omega)} \leq \left\| \nabla \left(\boldsymbol{v} - I_{\mathrm{div}}^h \boldsymbol{v}\right) \right\|_{L^2(\Omega)} \leq C h_K^m \left\| D^{m+1} \boldsymbol{v} \right\|_{\Omega}. \tag{3.71}$$

The constant C in (3.70) and (3.71) depends on the inverse of local discrete inf-sup constants. Operators I^h that satisfy the assumptions (3.67) and (3.68) are known, e.g, for space P_l and Q_l with $l \geq d$. However, the available theory does not apply for the Taylor-Hood spaces P_2/P_1 and Q_2/Q_1 in three dimensions. For concrete constructions of the operator I_{div}^h, it is referred to Girault and Scott (2003). □

Lemma 3.63 (Existence of an Operator I_{div}^h Implies Satisfaction of the Discrete Inf-Sup Condition) *Given an operator I_{div}^h with* (3.69) *and* (3.70)*, then the discrete inf-sup condition is satisfied with a constant independent of h, but depending on the inverse of the constant from estimate* (3.70)*.*

Proof Using (3.69), the triangle inequality, (3.70) for $k = 1$ and $m = 0$, and the inf-sup condition (3.48) for the continuous spaces yields

$$\begin{aligned}
&\inf_{q^h \in Q^h \setminus \{0\}} \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{\left(\nabla \cdot \boldsymbol{v}^h, q^h\right)}{\|\boldsymbol{v}^h\|_{V^h} \|q^h\|_{Q^h}} \\
&\geq \inf_{q^h \in Q^h \setminus \{0\}} \frac{\left(\nabla \cdot I_{\mathrm{div}}^h \boldsymbol{v}, q^h\right)}{\left\| I_{\mathrm{div}}^h \boldsymbol{v} \right\|_{V^h} \|q^h\|_{Q^h}} = \inf_{q^h \in Q^h \setminus \{0\}} \frac{\left(\nabla \cdot \boldsymbol{v}, q^h\right)}{\left\| I_{\mathrm{div}}^h \boldsymbol{v} \right\|_{V^h} \|q^h\|_{Q^h}}
\end{aligned}$$

$$\geq \inf_{q^h \in Q^h \setminus \{0\}} \frac{\left(\nabla \cdot \boldsymbol{v}, q^h\right)}{\left(\|\boldsymbol{v}\|_V + \left\|\boldsymbol{v} - I_{\text{div}}^h \boldsymbol{v}\right\|_{V^h}\right) \|q^h\|_{Q^h}}$$

$$\geq \inf_{q^h \in Q^h \setminus \{0\}} \frac{\left(\nabla \cdot \boldsymbol{v}, q^h\right)}{(1+C)\, \|\boldsymbol{v}\|_V \, \|q^h\|_{Q^h}} \geq \inf_{q \in Q \setminus \{0\}} \frac{(\nabla \cdot \boldsymbol{v}, q)}{(1+C)\, \|\boldsymbol{v}\|_V \, \|q\|_Q} \geq \frac{1}{(1+C)\beta_{\text{is}}},$$

where C is the constant from (3.70). ■

Remark 3.64 (Jumps Across Faces and Averages on Faces of Functions) Consider a triangulation $\mathcal{T}^h$ and let $K_1, K_2 \in \mathcal{T}^h$ be two mesh cells with a common $(d-1)$ face $E = K_1 \cap K_2$. Without loss of generality, the normal $\boldsymbol{n}_E$ at E should be the outward normal with respect to K_1. Then, the jump of a function v across the face E in the point $\boldsymbol{x} \in E$ is defined by

$$[|v|]_E = \lim_{\boldsymbol{y} \to \boldsymbol{x}, \boldsymbol{y} \in K_1} v(\boldsymbol{y}) - \lim_{\boldsymbol{y} \to \boldsymbol{x}, \boldsymbol{y} \in K_2} v(\boldsymbol{y}), \quad \boldsymbol{x} \in E, \tag{3.72}$$

if both limits are well defined. Changing the direction of $\boldsymbol{n}_E$ changes the sign of the jump.

The average is defined by

$$\{\{v\}\}_E = \frac{\lim_{\boldsymbol{y} \to \boldsymbol{x}, \boldsymbol{y} \in K_1} v(\boldsymbol{y}) + \lim_{\boldsymbol{y} \to \boldsymbol{x}, \boldsymbol{y} \in K_2} v(\boldsymbol{y})}{2}, \quad \boldsymbol{x} \in E.$$

Straightforward calculations, using these definitions, show

$$\begin{aligned} [|v+w|]_E &= [|v|]_E + [|w|]_E\,, \\ \{\{v+w\}\}_E &= \{\{v\}\}_E + \{\{w\}\}_E\,, \\ [|vw|]_E &= [|v|]_E \{\{w\}\}_E + \{\{v\}\}_E [|w|]_E\,. \end{aligned} \tag{3.73}$$

If w is continuous almost everywhere on E, then it follows from (3.73)

$$[|vw|]_E = [|v|]_E\, w.$$

□

Remark 3.65 (Sets of $(d-1)$ Faces) The set of all $(d-1)$ faces will be denoted by $\overline{\mathcal{E}}^h$ and the set of all faces which are not part of the boundary of Ω will be denoted by $\mathcal{E}^h$. □

Lemma 3.66 (Sufficient and Necessary Condition for a Finite Element Function to be in $H(\text{div}, \Omega)$, i.e., to Possess a Divergence in $L^2(\Omega)$) *Let $\mathcal{T}^h$ be a regular triangulation of Ω. A finite element function $\boldsymbol{v}^h \in L^2(\Omega)$, i.e., a piecewise polynomial function belongs to $H(\text{div}, \Omega)$, see (3.37), if and only if $\boldsymbol{v}^h \cdot \boldsymbol{n}_E$ is continuous for all faces E of the triangulation.*

Proof It has to be shown that $\nabla \cdot \boldsymbol{v}^h \in L^2(\Omega)$ if and only if the normal component of $\boldsymbol{v}^h$ is continuous for all faces. By definition, $\nabla \cdot \boldsymbol{v}^h \in L^2(\Omega)$ if and only if there exists a function $w \in L^2(\Omega)$ such that

$$-\int_\Omega \boldsymbol{v}^h(\boldsymbol{x}) \cdot \nabla\varphi(\boldsymbol{x})\, d\boldsymbol{x} = \int_\Omega w(\boldsymbol{x})\varphi(\boldsymbol{x})\, d\boldsymbol{x} \quad \forall\, \varphi \in C_0^\infty(\Omega). \tag{3.74}$$

Integration by parts yields

$$\begin{aligned}
&-\int_\Omega \boldsymbol{v}^h(\boldsymbol{x}) \cdot \nabla\varphi(\boldsymbol{x})\, d\boldsymbol{x} \\
&= -\sum_{K\in\mathcal{T}^h} \int_K \boldsymbol{v}^h(\boldsymbol{x}) \cdot \nabla\varphi(\boldsymbol{x})\, d\boldsymbol{x} \\
&= \sum_{K\in\mathcal{T}^h} \left(\int_K \nabla\cdot\boldsymbol{v}^h(\boldsymbol{x})\varphi(\boldsymbol{x})\, d\boldsymbol{x} - \int_{\partial K} \varphi(\boldsymbol{s})\boldsymbol{v}^h(\boldsymbol{s})\cdot\boldsymbol{n}_{\partial K}\, d\boldsymbol{s} \right) \\
&= \sum_{K\in\mathcal{T}^h} \int_K \nabla\cdot\boldsymbol{v}^h(\boldsymbol{x})\varphi(\boldsymbol{x})\, d\boldsymbol{x} - \sum_{K\in\mathcal{T}^h}\sum_{E\in\partial K} \int_E \varphi(\boldsymbol{s})\boldsymbol{v}^h(\boldsymbol{s})\cdot\boldsymbol{n}_E\, d\boldsymbol{s} \\
&= \int_\Omega \nabla\cdot\boldsymbol{v}^h(\boldsymbol{x})\varphi(\boldsymbol{x})\, d\boldsymbol{x} - \sum_{E\in\mathcal{E}^h} \int_E \varphi(\boldsymbol{s}) \left[\left|\boldsymbol{v}^h\cdot\boldsymbol{n}_E\right|\right]_E(\boldsymbol{s})\, d\boldsymbol{s} \\
&\quad - \sum_{E\in\overline{\mathcal{E}}^h\setminus\mathcal{E}^h} \int_E \varphi(\boldsymbol{s})\boldsymbol{v}^h(\boldsymbol{s})\cdot\boldsymbol{n}_E\, d\boldsymbol{s} \quad \forall\, \varphi\in C_0^\infty(\Omega).
\end{aligned} \tag{3.75}$$

The normal $\boldsymbol{n}_E$ on the interior faces can be chosen arbitrarily. Using the opposite normal $-\boldsymbol{n}_E$, also the sign of the jump has to be changed, i.e., one obtains

$$-\left[\left|\boldsymbol{v}^h\cdot(-\boldsymbol{n}_E)\right|\right]_E(\boldsymbol{s}) = \left[\left|\boldsymbol{v}^h\cdot\boldsymbol{n}_E\right|\right]_E(\boldsymbol{s}),$$

where (3.73) was applied. The last term in (3.75) vanishes since the test function vanishes at the boundary of Ω. Thus, (3.74) is satisfied if and only if all integrals on the interior faces vanish for all test functions. Therefore, the jumps $\left[\left|\boldsymbol{v}^h\cdot\boldsymbol{n}_E\right|\right]_E$ have to vanish on all interior faces, which is equivalent with the requirement that the normal component of $\boldsymbol{v}^h$ is continuous across all faces of the mesh cells. ■

3.4 Examples of Pairs of Finite Element Spaces Violating the Discrete Inf-Sup Condition

Remark 3.67 (On Low Order Pairs of Finite Element Spaces) The simplest and most common finite element spaces are spaces of continuous functions which are piecewise polynomials of first order, i.e., the P_1 space on triangular or tetrahedral grids and the Q_1 space on quadrilateral or hexahedral grids. Since for the pressure finite element space to be conforming it suffices that $Q^h \subset L_0^2(\Omega)$, one can use for Q^h also discontinuous finite element spaces, in the simplest case piecewise constants, i.e., P_0 or Q_0. However, the investigation of the discrete inf-sup condition (3.51) will show that this condition is not satisfied for pairs of those simple finite element spaces. These results are meanwhile classical and they can be found, e.g., in Girault and Raviart (1986, Chap. II, § 3.3) or Ern and Guermond (2004, Section 4.3.2). They have attracted a lot of research on appropriate finite element spaces for incompressible flow problems. □

Remark 3.68 (A Condition for the Violation of the Discrete Inf-Sup Condition) The violation of the discrete inf-sup condition (3.51) is proved, e.g., if one finds a non-trivial $q^h \in Q^h$ such that

$$b^h\left(\boldsymbol{v}^h, q^h\right) = 0 \quad \forall\, \boldsymbol{v}^h \in V^h. \tag{3.76}$$

In this case, it holds

$$\sup_{\boldsymbol{v}^h \in V^h, \boldsymbol{v}^h \neq \boldsymbol{0}} \frac{b^h\left(\boldsymbol{v}^h, q^h\right)}{\|\boldsymbol{v}^h\|_{V^h}} = 0,$$

from what follows, by dividing by $\left\|q^h\right\|_Q > 0$ and taking on both sides the infimum of the finite element pressure functions, that the discrete inf-sup condition (3.51) cannot be satisfied. Such non-trivial $q^h \in Q^h$ are called spurious pressure modes. □

Example 3.69 (The P_1/P_1 Pair of Finite Element Spaces) Probably every finite element code which uses simplicial grids can apply the P_1 finite element. If it would be possible to choose P_1/P_1 for velocity and pressure finite elements, the extension of such codes to the simulation of incompressible flows would be straightforward. However, this example shows that P_1/P_1 does not satisfy the discrete inf-sup condition (3.51).

Let $\Omega = (0, 1)^2$ and consider a triangulation of Ω with equally sized triangles with measure $|K| > 0$. Both, the finite element velocity and the finite element pressure are continuous and piecewise linear functions. The nodes of the finite element functions are their values in the vertices of the triangles, see Example B.38.

Consider first the integral mean value condition for the pressure, $Q^h \subset L_0^2(\Omega)$. Let K be a mesh cell and $q_{1,K}^h, q_{2,K}^h, q_{3,K}^h$ be the values of the pressure in the vertices of K. Then, the integral of q^h on K can be evaluated exactly by a quadrature rule which uses only the values at the vertices of K

$$\int_K q^h(\boldsymbol{x})\, d\boldsymbol{x} = \frac{|K|}{3}\left(q_{1,K}^h + q_{2,K}^h + q_{3,K}^h\right).$$

Hence, the integral mean value condition for the finite element pressure reads as follows

$$0 = \int_\Omega q^h(\boldsymbol{x})\, d\boldsymbol{x} = \sum_{K\in\mathcal{T}^h} \int_K q^h(\boldsymbol{x})\, d\boldsymbol{x} = \frac{|K|}{3} \sum_{K\in\mathcal{T}^h} \left(q_{1,K}^h + q_{2,K}^h + q_{3,K}^h\right). \tag{3.77}$$

Now, a function $q^h \in Q^h$ will be constructed that satisfies (3.76). On each mesh cell K, it is

$$\boldsymbol{v}^h\big|_K(\boldsymbol{x}) = \begin{pmatrix} \alpha_{11}x_1 + \alpha_{12}x_2 + \gamma_1 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \gamma_2 \end{pmatrix},$$

from what follows that

$$\nabla \cdot \boldsymbol{v}^h\big|_K(\boldsymbol{x}) = \alpha_{11} + \alpha_{22} = c_K.$$

Then, (3.76) becomes

$$\begin{aligned} 0 &= -\sum_{K\in\mathcal{T}^h} \int_K \left(\left(\nabla \cdot \boldsymbol{v}^h\right) q^h\right)(\boldsymbol{x})\, d\boldsymbol{x} = -\sum_{K\in\mathcal{T}^h} c_K \int_K q^h(\boldsymbol{x})\, d\boldsymbol{x} \\ &= -\frac{|K|}{3} \sum_{K\in\mathcal{T}^h} c_K \left(q_{1,K}^h + q_{2,K}^h + q_{3,K}^h\right). \end{aligned} \tag{3.78}$$

From (3.77) and (3.78), it follows that a counterexample for the fulfillment of the discrete inf-sup condition (3.51) is found, if a non-trivial function q^h with

$$q_{1,K}^h + q_{2,K}^h + q_{3,K}^h = 0$$

for all $K \in \mathcal{T}^h$ can be constructed. In this case, the integral mean value condition and (3.76) are satisfied both. Two examples of such functions are given in Fig. 3.1. The form of the spurious modes led to the name checkerboard-type instabilities. □

Example 3.70 (The P_1/P_0 Pair of Finite Element Spaces) As it was noted in Remark 3.52, the pressure finite element space should be sufficiently small for the fulfillment of the discrete inf-sup condition. From Example 3.69, it follows that this is obviously not the case for the P_1/P_1 finite element. On the first glance, a straightforward idea consists in approximating the pressure by even lower order polynomials, i.e., by piecewise constants instead of by continuous piecewise linears. The space P_0 consists of discontinuous functions, see Example B.37. Piecewise constant functions can be implemented without difficulties and the P_1/P_0 pair of finite element spaces would be easy to use, if it satisfies the discrete inf-sup condition (3.51).

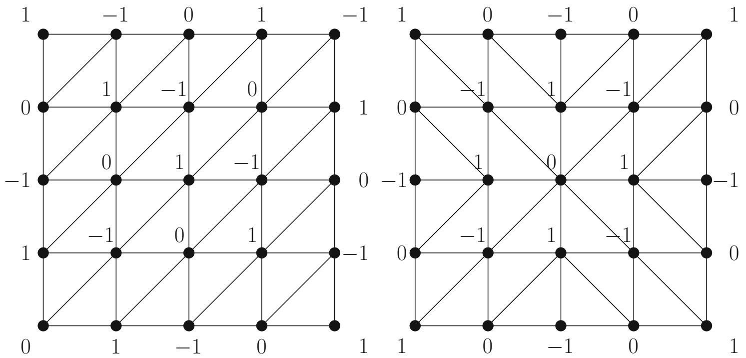

Fig. 3.1 Checkerboard instabilities for the P_1/P_1 finite element

For the pair P_1/P_0, it holds that $\nabla \cdot P_1 \subset P_0$. According to Remark 3.56, the discretely divergence-free finite element velocities have the desirable property to be even weakly divergence-free. However, the pair P_1/P_0 does not satisfy the discrete inf-sup condition (3.51).

From Fig. 3.1, it becomes clear that the dimension of the scalar-value space P_0 might be even larger than the dimension of the scalar-value space P_1. On these meshes, P_1 possesses 25 degrees of freedom (number of vertices) whereas P_0 has even 32 degrees of freedom (number of mesh cells). There is one degree of freedom less in the corresponding pressure finite element spaces since the discrete pressure has to fulfill the integral mean value condition. Thus, spurious modes have to be expected, which will be demonstrated now in detail. Besides this issue, another phenomenon can occur, which is called locking.

First, for showing that spurious modes can appear, consider for simplicity $\Omega = (0,1)^2$ and grids of the types presented in Fig. 3.1 with $2n^2$ mesh cells, $n > 1$. Then, the dimension of V^h corresponds to twice the number of interior vertices, $\dim\left(V^h\right) = 2\,(n-1)^2$. The dimension of Q^h corresponds to the number of mesh cells reduced by one (because of the integral mean value condition), i.e., $\dim\left(Q^h\right) = 2n^2 - 1$. It follows that

$$\dim\left(V^h\right) - \dim\left(Q^h\right) = 2n^2 - 4n + 2 - 2n^2 + 1 = 3 - 4n < 0$$

for $n \geq 1$. Hence, the necessary condition for the unique solvability of the discrete system, $\dim\left(V^h\right) \geq \dim\left(Q^h\right)$, see Remark 3.7, is not satisfied. Thus, there are at least $(4n-3)$ rows in B^h or columns in $\left(B^h\right)^T$ linearly dependent. It follows that there is a subspace of Q^h with dimension at least $(4n-3)$ such that for all q^h from this subspace it holds

$$0 = \left\langle \left(B^h\right)^T q^h, \boldsymbol{v}^h \right\rangle_{(V^h)',V^h} = b\left(\boldsymbol{v}^h, q^h\right) = -\left(\nabla \cdot \boldsymbol{v}^h, q^h\right) \quad \forall\ \boldsymbol{v}^h \in V^h,$$

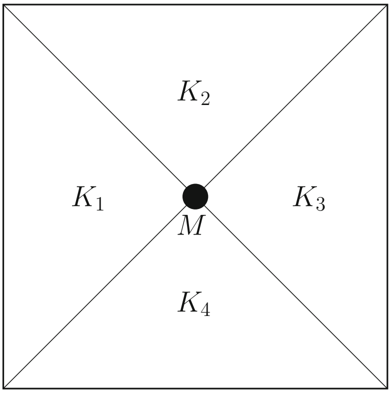

Fig. 3.2 Grid for the locking phenomenon of the P_1/P_0 finite element

which is (3.76). An extension of this example to general polygonal domains in two dimensions is possible, see Ern and Guermond (2004, p. 189).

Second, the locking phenomenon will be illustrated. Consider again $\Omega = (0,1)^2$ and a triangulation with four triangles that are obtained by drawing the two diagonals in Ω, see Fig. 3.2. Denote the piecewise constant pressure test function q^h restricted to the triangles K_i, $i = 1,\ldots,4$, by $q_1^h,\ldots,q_4^h$. Since all triangles have the same area, it follows from the mean value condition $q^h \in L_0^2(\Omega)$ that

$$q_1^h + q_2^h + q_3^h + q_4^h = 0.$$

Thus, there are three independent degrees of freedom for the discrete pressure. It follows that a function $\boldsymbol{v}^h \in V_{\text{div}}^h$ has to satisfy three independent conditions, see the definition (3.62) of this space. However, there are only two (scalar) degrees of freedom for the velocity, that are situated in the cross point M of the diagonals. Let $\boldsymbol{v}^h = (u^h, v^h)^T$, then one finds with a direct calculation that

$$\nabla\cdot\boldsymbol{v}^h\big|_{K_1} = -\nabla\cdot\boldsymbol{v}^h\big|_{K_3} = \frac{u^h(M)}{2}, \quad \nabla\cdot\boldsymbol{v}^h\big|_{K_4} = -\nabla\cdot\boldsymbol{v}^h\big|_{K_2} = \frac{v^h(M)}{2}.$$

Let, without loss of generality, q_1^h, q_2^h, q_3^h be the independent degrees of freedom for the discrete pressure. Then, one obtains from the condition for $\boldsymbol{v}^h$ being an element of V_{div}^h

$$\begin{aligned}
0 &= \left(\nabla\cdot\boldsymbol{v}^h, q^h\right) \\
&= |K_1|\left(q_1^h\,\nabla\cdot\boldsymbol{v}^h\big|_{K_1} + q_2^h\,\nabla\cdot\boldsymbol{v}^h\big|_{K_2} + q_3^h\,\nabla\cdot\boldsymbol{v}^h\big|_{K_3} - \left(q_1^h+q_2^h+q_3^h\right)\nabla\cdot\boldsymbol{v}^h\big|_{K_4}\right) \\
&= |K_1|\left(u^h(M)\left(\frac{q_1^h}{2}-\frac{q_3^h}{2}\right) - v^h(M)\left(\frac{q_1^h}{2}+q_2^h+\frac{q_3^h}{2}\right)\right),
\end{aligned}$$

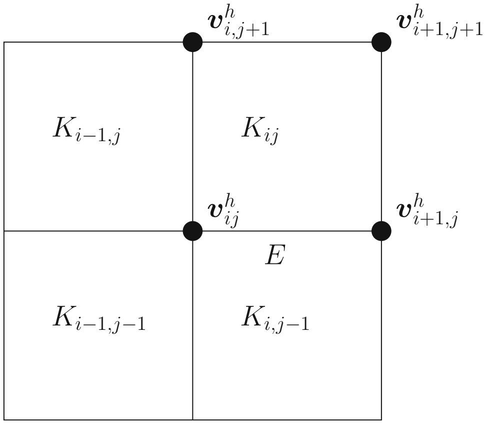

Fig. 3.3 Notation for Example 3.71

for arbitrary q_1^h, q_2^h, q_3^h. Hence $u^h(M) = v^h(M) = 0$ and $V_{\text{div}}^h = \{\mathbf{0}\}$. The situation that the only discretely divergence-free velocity is the trivial velocity is called locking phenomenon. □

Example 3.71 (The Q_1/Q_0 Pair of Finite Element Spaces) Consider $\Omega = (0,1)^2$ and a triangulation of Ω that consists of $n \times n$ squares, where n is even. The squares are denoted by K_{ij}, $i, j = 1, \ldots, n$. A lexicographic enumeration (left to right, bottom to top) is used. The finite element velocity is piecewise bilinear and it is determined by its values at the vertices of the squares K_{ij}, see Example B.50. Denote the value at the lower left vertex by $\boldsymbol{v}_{ij}^h = \left(u_{ij}^h, v_{ij}^h\right)^T$, see Fig. 3.3. A piecewise constant finite element function is used for approximating the pressure, see Example B.49. Its value in K_{ij} is denoted by q_{ij}^h.

Compared with the P_1/P_0 finite element on a corresponding grid, where the squares are divided by their diagonals, the number of degrees of freedom for the finite element velocity is the same, since the number of interior vertices is equal. However, the number of degrees of freedom of the Q_0 finite element pressure space (number of mesh cells minus one) is only roughly half of the corresponding P_0 finite element space. Following Remark 3.52, the decrease of the dimension of Q^h might be a way to obtain an inf-sup stable pair of finite element spaces.

Since the integral mean value of finite element pressures should be zero, it follows that

$$\int_\Omega q^h(\boldsymbol{x})\, d\boldsymbol{x} = \sum_{i,j=1}^{n} \int_{K_{ij}} q_{ij}^h \, d\boldsymbol{x} = h^2 \sum_{i,j=1}^{n} q_{ij}^h = 0,$$

where h is the length of the edges of K_{ij} and $h^2 = \left|K_{ij}\right|$. Hence, it holds

$$\sum_{i,j=1}^{n} q_{ij}^h = 0. \tag{3.79}$$

Consider a single mesh cell K_{ij}. Integration by parts gives

$$\int_{K_{ij}} \left(\nabla\cdot\boldsymbol{v}^h\right) q_{ij}^h \, d\boldsymbol{x} = q_{ij}^h \int_{K_{ij}} \nabla\cdot\boldsymbol{v}^h \, d\boldsymbol{x} = q_{ij}^h \int_{\partial K_{ij}} \boldsymbol{v}^h\cdot\boldsymbol{n}_{\partial K_{ij}} \, ds,$$

where $\boldsymbol{v}^h = (u^h, v^h)^T$ and $\boldsymbol{n}_{\partial K_{ij}}$ denotes the outward pointing unit normal on ∂K_{ij}. The outward pointing normals on ∂K_{ij} are $\pm\boldsymbol{e}_i$, $i = 1, 2$, and the restriction of a Q_1 finite element function to an edge is a linear function. That means, the integrals on the edges can be computed exactly by the trapezoidal rule. One obtains, e.g., for the bottom edge E

$$\int_E \boldsymbol{v}^h\cdot\boldsymbol{n}_{\partial K_{ij}} \, ds = -\int_E \boldsymbol{v}^h\cdot\boldsymbol{e}_2 \, ds = -\int_E v^h(s) \, ds = -\frac{h}{2}\left(v_{ij}^h + v_{i+1,j}^h\right).$$

Altogether, it is

$$\begin{aligned}
&\int_{K_{ij}} \left(\left(\nabla\cdot\boldsymbol{v}^h\right) q^h\right)(\boldsymbol{x}) \, d\boldsymbol{x} = \left(\nabla\cdot\boldsymbol{v}^h, q^h\right)_{K_{ij}} \\
&= \frac{h}{2} q_{ij}^h \Big(-v_{ij}^h - v_{i+1,j}^h + u_{i+1,j}^h + u_{i+1,j+1}^h + v_{i+1,j+1}^h + v_{i,j+1}^h - u_{i,j+1}^h - u_{ij}^h \Big).
\end{aligned}$$

Now, the contributions of the integrals that are connected to an interior velocity node are considered. Nodes on the boundary are not of interest since the velocity vanishes in these nodes. One gets for the node ij, $i,j = 2,\ldots,n$, the following contributions from the four mesh cells for which this node is a vertex

$$\begin{aligned}
&u_{ij}^h \frac{h}{2}\left(q_{i-1,j-1}^h + q_{i-1,j}^h - q_{i,j-1}^h - q_{ij}^h\right), \\
&v_{ij}^h \frac{h}{2}\left(q_{i-1,j-1}^h + q_{i,j-1}^h - q_{i-1,j}^h - q_{ij}^h\right).
\end{aligned}$$

The sum of the integrals on the mesh cells is the same as the sum of the contributions connected to the interior velocity nodes, because this is just a rearrangement of the sum. Hence, condition (3.76) can be written in the form

$$\begin{aligned}
0 &= \left(\nabla\cdot\boldsymbol{v}^h, q^h\right) = \sum_{i,j=1}^{n} \left(\nabla\cdot\boldsymbol{v}^h, q^h\right)_{K_{ij}} \\
&= \frac{h}{2}\sum_{i,j=2}^{n} \Big[u_{ij}^h \left(q_{i-1,j-1}^h + q_{i-1,j}^h - q_{i,j-1}^h - q_{ij}^h\right) \\
&\quad + v_{ij}^h \left(q_{i-1,j-1}^h + q_{i,j-1}^h - q_{i-1,j}^h - q_{ij}^h\right) \Big] \quad \forall\ \boldsymbol{v}^h \in V^h.
\end{aligned}$$

In particular, this equality holds for all finite element velocities where just one value u_{ij}^h or v_{ij}^h is non-zero. Inserting these finite element functions, it follows that the finite element pressure has to fulfill

$$q_{i-1,j-1}^h + q_{i-1,j}^h - q_{i,j-1}^h - q_{ij}^h = 0, \quad q_{i-1,j-1}^h + q_{i,j-1}^h - q_{i-1,j}^h - q_{ij}^h = 0,$$

$i,j = 2,\ldots n$. Adding these conditions, respectively subtracting them, leads to the new conditions

$$q_{ij}^h = q_{i-1,j-1}^h, \quad q_{i,j-1}^h = q_{i-1,j}^h \quad i,j = 2,\ldots n. \tag{3.80}$$

With (3.80), there are on the first glance two free choices for setting the finite element pressure values. Setting on a cell, e.g., on K_{11} the value $\alpha \in \mathbb{R}$, then half of the mesh cells, i.e., $n^2/2$ mesh cells, get the same value. These are all mesh cells where the sum of the indices is even. Setting on another mesh cell, where the sum of the indices is odd, the value β, then the second half of the mesh cells get this value. For dividing the set of mesh cells into two subsets with the same number of elements, the assumption that n is even is needed. From the condition on the vanishing integral mean value (3.79), one gets

$$0 = \sum_{i,j=1, i+j \text{ even}}^{n} \alpha + \sum_{i,j=1, i+j \text{ odd}}^{n} \beta = \frac{n^2}{2}(\alpha + \beta).$$

Hence, it is $\beta = -\alpha$, $\alpha \in \mathbb{R}$. In summary, for all $\alpha \neq 0$, a non-trivial finite element pressure was found that satisfies (3.76), for an example see Fig. 3.4. Because of its particular form, this instability is called checkerboard instability. □

Fig. 3.4 A checkerboard instability for the Q_1/Q_0 finite element

−1	1	−1	1
1	−1	1	−1
−1	1	−1	1
1	−1	1	−1

Remark 3.72 (Modifications of the Q_1/Q_0 Finite Element for Improving the Stability)

- Since the spurious mode can be well characterized, one could try to define a smaller finite element pressure space by taking only those functions from Q^h which are $L^2(\Omega)$ orthogonal to this mode. It was shown that this approach leads to an inf-sup stable pair of finite element spaces. However, the inf-sup parameter depends on the mesh width, $\beta_{\mathrm{is}}^h = Ch$, with C independent of h, e.g., see Girault and Raviart (1986, II, §3). It will become obvious in the finite element error analysis that the inverse of β^h enters the estimates. Thus, this method might not converge.
- In Boland and Nicolaides (1985), a further modification of the pressure finite element space was proposed that leads to a uniformly in h stable pair of finite element spaces, see also Girault and Raviart (1986, p. 167) for details. However, to the best of our knowledge, the method from Boland and Nicolaides (1985) is not used in practice.
- The fulfillment of the discrete inf-sup condition for the Q_1/Q_0 pair of finite element spaces can be proved on a very special mesh, see Stenberg (1984).

□

Example 3.73 ($P_k/P_{k-1}^{\mathrm{disc}}$, $k \geq 2$, on a Special Macro Cell) Consider the pair of spaces $P_k/P_{k-1}^{\mathrm{disc}}$, $k \geq 2$, on the grid shown in Fig. 3.5, in particular in the macro cell which is surrounded boldly. The diagonals of the macro cell should be parallel to the coordinate axes. The pair $P_k/P_{k-1}^{\mathrm{disc}}$, $k \geq 2$, is called Scott–Vogelius finite element, see Scott and Vogelius (1985).

Let $\boldsymbol{v}^h \in P_k$, then $\nabla \cdot \boldsymbol{v}^h \in P_{k-1}^{\mathrm{disc}} = Q^h$. Hence, Remark 3.56 gives that discretely divergence-free finite element velocities from $P_k/P_{k-1}^{\mathrm{disc}}$ are even weakly divergence-free, which is a desirable property. However, it turns out that $P_k/P_{k-1}^{\mathrm{disc}}$ does not fulfill the discrete inf-sup condition on the mesh presented in Fig. 3.5.

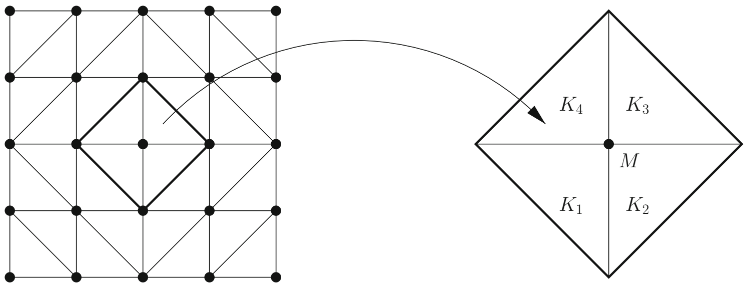

Fig. 3.5 Grid for $P_k/P_{k-1}^{\mathrm{disc}}$, $k \geq 2$, finite element

Consider $\boldsymbol{v}^h = (v_1^h, v_2^h)^T \in V_{\text{div}}^h$. Since $\boldsymbol{v}^h$ is continuous, the tangential derivatives $\nabla \boldsymbol{v}^h \cdot \mathbf{t}$ on the edges are continuous, too. Using that the tangentials are Cartesian unit vectors, one obtains

$$\frac{\partial v_2^h|_{K_1}}{\partial y} = \frac{\partial v_2^h|_{K_2}}{\partial y}, \quad \frac{\partial v_1^h|_{K_2}}{\partial x} = \frac{\partial v_1^h|_{K_3}}{\partial x},$$
$$\frac{\partial v_2^h|_{K_3}}{\partial y} = \frac{\partial v_2^h|_{K_4}}{\partial y}, \quad \frac{\partial v_1^h|_{K_4}}{\partial x} = \frac{\partial v_1^h|_{K_2}}{\partial x}.$$

These equations hold in particular in the vertex M. Using in addition the pointwise divergence-free constraint in K_2, K_3, and K_4, it follows that

$$\begin{aligned}
&\frac{\partial v_1^h|_{K_1}}{\partial x}(M) + \frac{\partial v_2^h|_{K_1}}{\partial y}(M) \\
&= \frac{\partial v_1^h|_{K_1}}{\partial x}(M) + \frac{\partial v_2^h|_{K_2}}{\partial y}(M) = \frac{\partial v_1^h|_{K_1}}{\partial x}(M) - \frac{\partial v_1^h|_{K_2}}{\partial x}(M) \\
&= \frac{\partial v_1^h|_{K_1}}{\partial x}(M) - \frac{\partial v_1^h|_{K_3}}{\partial x}(M) = \frac{\partial v_1^h|_{K_1}}{\partial x}(M) + \frac{\partial v_2^h|_{K_3}}{\partial y}(M) \\
&= \frac{\partial v_1^h|_{K_1}}{\partial x}(M) + \frac{\partial v_2^h|_{K_4}}{\partial y}(M) = \frac{\partial v_1^h|_{K_1}}{\partial x}(M) - \frac{\partial v_1^h|_{K_4}}{\partial x}(M) \\
&= \frac{\partial v_1^h|_{K_1}}{\partial x}(M) - \frac{\partial v_1^h|_{K_1}}{\partial x}(M) = 0.
\end{aligned}$$

That means, the divergence-free condition in M for the restriction of $\boldsymbol{v}^h$ to K_1 follows from the continuity of the tangential derivatives on the edges and the divergence-free conditions in K_2, K_3, and K_4.

Since $\boldsymbol{v}^h$ is weakly divergence-free, it is

$$\left(\nabla \cdot \boldsymbol{v}^h, q^h\right)_{K_1} = 0 \quad \forall\, q^h \in Q^h(K_1). \tag{3.81}$$

One deduces from (3.81) that there are $\dim P_{k-1}(K_1) = \dim Q^h(K_1) = k(k+1)/2$ conditions to determine $\nabla \cdot \boldsymbol{v}^h \in P_{k-1}(K_1)$ and there is independently the derived condition $\nabla \cdot \boldsymbol{v}^h\big|_{K_1}(M) = 0$. Hence, there is one condition too much in (3.81). Since one of the conditions from (3.81) can be neglected, it follows that one of the pressure degrees of freedom in K_1 is not determined by (3.81). This pressure degree of freedom becomes a spurious node. □

Remark 3.74 (The Scott–Vogelius Element on Other Meshes) The consideration of the special grid in Example 3.73 is crucial for proving that the Scott–Vogelius finite element does not satisfy the discrete inf-sup condition. It can be shown that Scott–Vogelius pairs of finite element spaces of sufficiently high order fulfill the discrete inf-sup condition on meshes that are in some sense less structured than the mesh from Fig. 3.5, see Sect. 3.6.3. □

Remark 3.75 (Circumventing the Discrete Inf-Sup Condition) As mentioned in Remark 3.67, low order finite element spaces are quite popular. In order to use them for incompressible flow simulations, several proposals were presented that circumvent the need of satisfying the discrete inf-sup condition (3.51). The principal idea consists in removing the saddle point character of the problem by introducing a bilinear form in the continuity equation that leads to a pressure-pressure coupling. One example is the PSPG method discussed in Sect. 5.3. Of course, a consequence of introducing an additional term into the continuity equation results in a violation of the (discrete) mass balance. □

3.5 Techniques for Checking the Discrete Inf-Sup Condition

Remark 3.76 (Contents) This section presents some ways for checking the discrete inf-sup condition (3.51) or equivalently (3.52). A few applications of these ideas to concrete pairs of finite element spaces will be presented in Sect. 3.6.

A comprehensive overview of techniques for proving the discrete inf-sup condition and corresponding results can be found, e.g., in Boffi et al. (2008) and Boffi et al. (2013, Sections 8.4 and 8.5). □

3.5.1 The Fortin Operator

Remark 3.77 (A Connection Between the Continuous and the Discrete Inf-Sup Condition) For conforming finite element spaces, it is possible to check the discrete inf-sup condition (3.51) with the help of the continuous inf-sup condition (3.48). The connection of both conditions is shown in the following lemma. The result is due to Fortin (1977) and it can be generalized to Banach spaces, see Ern and Guermond (2004, Lemma 4.19). Also the generalization to some non-conforming cases is possible, see Sect. 3.6.5.

For conforming finite element spaces, it is $b^h(\cdot,\cdot) = b(\cdot,\cdot)$. □

Lemma 3.78 (Fortin Criterion for Checking the Discrete Inf-Sup Condition) *Let V, Q, and $b(\cdot,\cdot)$ fulfill the assumptions of Remark 3.3 and let the inf-sup condition (3.48) be satisfied. Consider conforming spaces $V^h \subset V$ and $Q^h \subset Q$. Then, V^h and Q^h satisfy the discrete inf-sup condition (3.51) if and only if there exists a constant $\gamma^h > 0$, which is independent of h, such that for all $\boldsymbol{v} \in V$ there is an element $P_{\mathrm{For}}^h \boldsymbol{v} \in V^h$ with*

$$b\left(\boldsymbol{v}, q^h\right) = b\left(P_{\mathrm{For}}^h \boldsymbol{v}, q^h\right) \quad \forall\, q^h \in Q^h \quad \textit{and} \quad \left\|P_{\mathrm{For}}^h \boldsymbol{v}\right\|_V \le \gamma^h \left\|\boldsymbol{v}\right\|_V . \tag{3.82}$$

Proof • *Assume that* (3.82) *holds.*

Let $q^h \in Q^h$ be arbitrary. From span $\left\{P_{\text{For}}^h \boldsymbol{v}\right\} \subseteq V^h$, it follows, using also (3.82) and (3.48), that

$$\sup_{\boldsymbol{v}^h \in V^h \setminus \{\mathbf{0}\}} \frac{b\left(\boldsymbol{v}^h, q^h\right)}{\left\|\boldsymbol{v}^h\right\|_V} \geq \sup_{\boldsymbol{v} \in V \setminus \{\mathbf{0}\}} \frac{b\left(P_{\text{For}}^h \boldsymbol{v}, q^h\right)}{\left\|P_{\text{For}}^h \boldsymbol{v}\right\|_V} = \sup_{\boldsymbol{v} \in V \setminus \{\mathbf{0}\}} \frac{b\left(\boldsymbol{v}, q^h\right)}{\left\|P_{\text{For}}^h \boldsymbol{v}\right\|_V}$$
$$\geq \sup_{\boldsymbol{v} \in V \setminus \{\mathbf{0}\}} \frac{b\left(\boldsymbol{v}, q^h\right)}{\gamma^h \left\|\boldsymbol{v}\right\|_V} \geq \frac{\beta_{\text{is}}}{\gamma^h} \left\|q^h\right\|_Q .$$

This inequality is just the discrete inf-sup condition (3.51) with $\beta_{\text{is}}^h = \beta_{\text{is}}/\gamma^h$.

- *Assume that* (3.51) *is satisfied.*

Consider the restriction of $b(\cdot,\cdot)$ from $V \times Q$ to $V \times Q^h$. This restriction defines a continuous linear operator

$$\tilde{B} \in \mathcal{L}\left(V, \left(Q^h\right)'\right), \quad \left\langle \tilde{B}\boldsymbol{v}, q^h \right\rangle_{(Q^h)',Q^h} = b\left(\boldsymbol{v}, q^h\right).$$

By definition, it is $\tilde{B}\boldsymbol{v} \in \left(Q^h\right)'$ for all $\boldsymbol{v} \in V$. Since the discrete inf-sup condition (3.51) holds, it follows from Lemma 3.12 iii) that the operator B^h defined in (3.61), the discrete divergence operator, is an isomorphism from $(V_{\text{div}}^h)^\perp$ onto $\left(Q^h\right)'$. In particular, B^h is surjective. Since $\tilde{B}\boldsymbol{v} \in \left(Q^h\right)'$, there must be an element $\tilde{\boldsymbol{v}}^h$ from $(V_{\text{div}}^h)^\perp$ such that

$$\left\langle B^h \tilde{\boldsymbol{v}}^h, q^h \right\rangle_{(Q^h)',Q^h} = \left\langle \tilde{B}\boldsymbol{v}, q^h \right\rangle_{(Q^h)',Q^h} \quad \forall\, q^h \in Q^h.$$

Consequently, for all $\boldsymbol{v}^h \in V^h$ whose projection into $(V_{\text{div}}^h)^\perp$ is equal to $\tilde{\boldsymbol{v}}^h$, it holds

$$\left\langle B^h \boldsymbol{v}^h, q^h \right\rangle_{(Q^h)',Q^h} = \left\langle \tilde{B}\boldsymbol{v}, q^h \right\rangle_{(Q^h)',Q^h} \quad \forall\, q^h \in Q^h. \tag{3.83}$$

One of these elements can be chosen to be $P_{\text{For}}^h \boldsymbol{v}$. Then, (3.83) is equivalent to

$$b\left(P_{\text{For}}^h \boldsymbol{v}, q^h\right) = b\left(\boldsymbol{v}, q^h\right) \quad \forall\, q^h \in Q^h.$$

With these relations, it follows from Lemma 3.12 iii), the definition of the norm in Q^h, the definition of the norm of $b(\cdot,\cdot)$ from (3.3), and the estimate of this norm from Lemma 3.39 that

$$\left\|P_{\text{For}}^h \boldsymbol{v}\right\|_V \leq \frac{1}{\beta_{\text{is}}^h} \left\|B^h\left(P_{\text{For}}^h \boldsymbol{v}\right)\right\|_Q = \frac{1}{\beta_{\text{is}}^h} \sup_{q^h \in Q^h \setminus \{0\}} \frac{\left\langle B^h\left(P_{\text{For}}^h \boldsymbol{v}\right), q^h \right\rangle_{(Q^h)',Q^h}}{\|q^h\|_Q}$$
$$= \frac{1}{\beta_{\text{is}}^h} \sup_{q^h \in Q^h \setminus \{0\}} \frac{b\left(P_{\text{For}}^h \boldsymbol{v}, q^h\right)}{\|q^h\|_Q} = \frac{1}{\beta_{\text{is}}^h} \sup_{q^h \in Q^h \setminus \{0\}} \frac{b(\boldsymbol{v}, q^h)}{\|q^h\|_Q}$$
$$\leq \frac{1}{\beta_{\text{is}}^h} \sup_{q^h \in Q^h \setminus \{0\}} \frac{\|b\| \, \|\boldsymbol{v}\|_V \left\|q^h\right\|_Q}{\|q^h\|_Q} = \frac{\|b\|}{\beta_{\text{is}}^h} \|\boldsymbol{v}\|_V = \gamma^h \|\boldsymbol{v}\|_V .$$

Since $\boldsymbol{v}$ was chosen to be arbitrary, (3.82) is proved. ■

Remark 3.79 (On Condition (3.76)) This condition, which implies that the discrete inf-sup condition is violated, cannot be fulfilled if (3.82) holds. Assume that there is a $q^h \in Q^h$ such that $b(\boldsymbol{v}^h, q^h) = 0$ for all $\boldsymbol{v}^h \in V^h$. From (3.82) it follows that then $b(\boldsymbol{v}, q^h) = 0$ for all $\boldsymbol{v} \in V$, since $P_{\mathrm{For}}^h \boldsymbol{v} \in V^h$. Because $q^h \in Q$ and V and Q satisfy the inf-sup condition (3.48), it follows that $q^h = 0$. Hence, there is no non-trivial $q^h \in Q^h$ for which (3.76) holds. □

Remark 3.80 (Kernel of the Gradient Operators) From condition (3.82), first part, one obtains $\ker(\mathrm{grad}^h) \subseteq \ker(\mathrm{grad})$. To show this inclusion, let $q^h \in \ker(\mathrm{grad}^h)$, then $b(\boldsymbol{v}^h, q^h) = 0$ for all $\boldsymbol{v}^h \in V^h$. Since $P_{\mathrm{For}}^h \boldsymbol{v} \in V^h$ for all $\boldsymbol{v} \in V$, it follows that $b(\boldsymbol{v}, q^h) = 0$ for all $\boldsymbol{v} \in V$. Hence, $q^h \in \ker(\mathrm{grad})$. □

Remark 3.81 (A Possible Construction of a Fortin Operator) Sometimes, it is possible to construct a linear Fortin operator P_{For}^h with the help of two linear operators $P_1^h, P_2^h \in \mathcal{L}(V, V^h)$. Assume that

$$\left\| P_1^h \boldsymbol{v} \right\|_V \le C_1 \left\| \boldsymbol{v} \right\|_V \quad \forall\, \boldsymbol{v} \in V, \tag{3.84}$$

$$\left\| P_2^h \left(I - P_1^h\right) \boldsymbol{v} \right\|_V \le C_2 \left\| \boldsymbol{v} \right\|_V \quad \forall\, \boldsymbol{v} \in V, \tag{3.85}$$

$$b\left(\boldsymbol{v} - P_2^h \boldsymbol{v}, q^h\right) = 0 \quad \forall\, \boldsymbol{v} \in V,\ \forall\, q^h \in Q^h, \tag{3.86}$$

where C_1, C_2 are independent of h. Then, a Fortin operator is defined by

$$P_{\mathrm{For}}^h \in \mathcal{L}(V, V^h) \quad \boldsymbol{v} \mapsto P_1^h \boldsymbol{v} + P_2^h \left(\boldsymbol{v} - P_1^h \boldsymbol{v}\right). \tag{3.87}$$

□

Lemma 3.82 (A Property of the Operator (3.87)) *The operator defined in (3.87) satisfies (3.82).*

Proof Applying (3.87), (3.86) for $P_1^h \boldsymbol{v}$, and once again (3.86) for $\boldsymbol{v}$, one obtains for all $q^h \in Q^h$

$$\begin{aligned} b\left(P_{\mathrm{For}}^h \boldsymbol{v}, q^h\right) &= b\left(P_1^h \boldsymbol{v} + P_2^h \left(\boldsymbol{v} - P_1^h \boldsymbol{v}\right), q^h\right) \\ &= b\left(P_1^h \boldsymbol{v} - P_2^h P_1^h \boldsymbol{v}, q^h\right) + b\left(P_2^h \boldsymbol{v}, q^h\right) \\ &= b\left(P_2^h \boldsymbol{v}, q^h\right) = b\left(\boldsymbol{v}, q^h\right). \end{aligned}$$

The boundedness of P_{For}^h is obtained by applying the triangle inequality and using (3.84) and (3.85)

$$\left\| P_{\mathrm{For}}^h \boldsymbol{v} \right\|_V \le \left\| P_1 \boldsymbol{v} \right\|_V + \left\| P_2^h \left(\boldsymbol{v} - P_1^h \boldsymbol{v}\right) \right\|_V \le C_1 \left\| \boldsymbol{v} \right\|_V + C_2 \left\| \boldsymbol{v} \right\|_V = \gamma^h \left\| \boldsymbol{v} \right\|_V,$$

with $\gamma^h = C_1 + C_2$. ■

Remark 3.83 (A More Detailed Construction of the Fortin Operator) Often, the Clément operator P^h_{Cle} (C.17), with the modification that preserves homogeneous Dirichlet boundary conditions, see Remark C.22, plays the role of P^h_1. Then, condition (3.85) for P^h_2 can be replaced with

$$\left\|P^h_2 \boldsymbol{v}\right\|_{H^1(K)} \le C\left(h_K^{-1}\left\|\boldsymbol{v}\right\|_{L^2(K)} + \left|\boldsymbol{v}\right|_{H^1(K)}\right), \quad \forall\, K \in \mathcal{T}^h,\ \forall\, \boldsymbol{v} \in V, \tag{3.88}$$

where the constant C does not depend on h_K. □

Lemma 3.84 (A Property of the Fortin Operator Constructed with (C.17), (3.86), and (3.88)) *Consider a family of quasi-uniform triangulations $\{\mathcal{T}^h\}$. Let $P^h_1 = P^h_{\mathrm{Cle}}$ be the modified Clément interpolation operator (C.17), which preserves homogeneous Dirichlet boundary conditions, and let P^h_2 satisfy (3.86) and (3.88). Then, P^h_{For} defined by (3.87) is a Fortin operator.*

Proof The first property of (3.82) is proved analogously as in the proof of Lemma 3.82, since the proof used only (3.86) and (3.87). It remains to show the second property with γ^h independent of h.

From the quasi-uniformity of the family of triangulations, it follows that for each K there is a maximal number of mesh cells in ω_K, see Fig. C.1, which is independent of the triangulation and that the diameter of ω_K can be estimated by Ch_K with a constant C independent of $\mathcal{T}^h$. Using (3.87), the triangle inequality, (3.88), and (C.18) for $k = 0, l = 1$ and $k = l = 1$, one obtains

$$\begin{aligned}
\left\|P^h_{\mathrm{For}}\boldsymbol{v}\right\|_V^2 &= \left|P^h_{\mathrm{For}}\boldsymbol{v}\right|^2_{H^1(\Omega)} \\
&\le 2\left|P^h_{\mathrm{Cle}}\boldsymbol{v}\right|^2_{H^1(\Omega)} + 2\left|P^h_2\left(\boldsymbol{v} - P^h_{\mathrm{Cle}}\boldsymbol{v}\right)\right|^2_{H^1(\Omega)} \\
&\le C\left(\left|P^h_{\mathrm{Cle}}\boldsymbol{v} - \boldsymbol{v}\right|^2_{H^1(\Omega)} + \left|\boldsymbol{v}\right|^2_{H^1(\Omega)}\right) + 2\sum_{K\in\mathcal{T}^h}\left|P^h_2\left(\boldsymbol{v} - P^h_{\mathrm{Cle}}\boldsymbol{v}\right)\right|^2_{H^1(K)} \\
&\le C\left(\sum_{K\in\mathcal{T}^h}\left|P^h_{\mathrm{Cle}}\boldsymbol{v} - \boldsymbol{v}\right|^2_{H^1(K)} + \left|\boldsymbol{v}\right|^2_{H^1(\Omega)}\right) \\
&\quad + C\left(\sum_{K\in\mathcal{T}^h}(h_K^{-2}\left\|\boldsymbol{v} - P^h_{\mathrm{Cle}}\boldsymbol{v}\right\|^2_{L^2(K)} + \left|\boldsymbol{v} - P^h_{\mathrm{Cle}}\boldsymbol{v}\right|^2_{H^1(K)}\right) \\
&\le C\left(\left|\boldsymbol{v}\right|^2_{H^1(\Omega)} + \sum_{K\in\mathcal{T}^h}\left(\left|\boldsymbol{v}\right|^2_{H^1(\omega_K)} + \left|\boldsymbol{v}\right|^2_{H^1(\omega_K)} + \left|\boldsymbol{v}\right|^2_{H^1(\omega_K)}\right)\right) \\
&\le C\left\|\nabla\boldsymbol{v}\right\|^2_{L^2(\Omega)} = C\left\|\boldsymbol{v}\right\|^2_V .
\end{aligned}$$

■

Remark 3.85 (Non-conforming Discretizations) The concept of the Fortin operator is also applied in non-conforming methods for proving the satisfaction of the discrete inf-sup condition, e.g., see the proof of Theorem 3.151 for the Crouzeix–Raviart pair of finite element spaces: Equations (3.144) and (3.145). □

3.5.2 Splitting the Discrete Pressure into a Piecewise Constant Part and a Remainder

Remark 3.86 (The Approach) It was noted in Remark 3.52 that a small pressure finite element space is advantageous to satisfy the discrete inf-sup condition (3.51) or equivalently (3.52). A small finite element space on a given triangulation is the space of piecewise constants which is intersected with $L_0^2(\Omega)$. The approach presented in this section uses the assumption that the discrete inf-sup condition holds for V^h/P_0 or V^h/Q_0. The finite element pressure on the right-hand side of discrete inf-sup condition (3.52) will be split into a piecewise constant part $\overline{q}^h$ and a remainder

$$q^h = \overline{q}^h + \left(q^h - \overline{q}^h\right).$$

For this remainder, also an inf-sup condition is assumed to prove finally the desired discrete inf-sup condition. The idea is that the proof of the inf-sup condition for both parts, the piecewise constant one and the remainder, is comparatively easy.

This approach was introduced in Brezzi and Bathe (1990). □

Lemma 3.87 (A Property of the $L^2(\Omega)$ Projection into a Piecewise Constant Finite Element Space) *Let $q^h \in Q^h$ and denote by $\overline{q}^h$ its $L^2(\Omega)$ projection into P_0 or Q_0*

$$\left(q^h - \overline{q}^h, \overline{r}^h\right) = 0 \quad \forall\, \overline{r}^h \in P_0(or\ Q_0). \tag{3.89}$$

Then

$$\overline{q}^h\big|_K = \frac{1}{|K|}\int_K q^h(\boldsymbol{x})\,d\boldsymbol{x} \quad \forall\, K \in \mathcal{T}^h,$$

where $|K|$ is the measure of K, and $\overline{q}^h \in L_0^2(\Omega)$.

Proof Using that $\overline{q}^h$ and $\overline{r}^h$ are piecewise constant, one obtains

$$\left(q^h, \overline{r}^h\right) = \sum_{K\in\mathcal{T}^h}\left(q^h, \overline{r}^h\right)_K = \sum_{K\in\mathcal{T}^h}\overline{r}^h\big|_K\int_K q^h(\boldsymbol{x})\,d\boldsymbol{x}$$

and

$$\left(\overline{q}^h, \overline{r}^h\right) = \sum_{K\in\mathcal{T}^h}\overline{r}^h\big|_K\,\overline{q}^h\big|_K\int_K d\boldsymbol{x} = \sum_{K\in\mathcal{T}^h}\overline{r}^h\big|_K\,\overline{q}^h\big|_K\,|K|\,.$$

Inserting both expressions in (3.89) proves the formula for $\overline{q}^h\big|_K$. Using this formula and $q^h \in L_0^2(\Omega)$ shows that the mean value of $\overline{q}^h$ vanishes, since

$$\begin{aligned}\int_\Omega \overline{q}^h(\boldsymbol{x})\,d\boldsymbol{x} &= \sum_{K\in\mathcal{T}^h}\int_K \overline{q}^h(\boldsymbol{x})\,d\boldsymbol{x} = \sum_{K\in\mathcal{T}^h} \overline{q}^h\big|_K \int_K d\boldsymbol{x}\\ &= \sum_{K\in\mathcal{T}^h}\frac{1}{|K|}\int_K q^h(\boldsymbol{x})\,d\boldsymbol{x}\int_K d\boldsymbol{x} = \int_\Omega q^h(\boldsymbol{x})\,d\boldsymbol{x} = 0.\end{aligned}$$

■

Remark 3.88 (A Modified Inf-Sup Condition) A central role in this approach plays an inf-sup condition of the form

$$\sup_{\boldsymbol{v}^h\in V^h\setminus\{\boldsymbol{0}\}} \frac{b^h(\boldsymbol{v}^h,q^h)}{\|\boldsymbol{v}^h\|_{V^h}} \ge \beta_1 \left\|q^h-\overline{q}^h\right\|_Q \quad \forall\, q^h\in Q^h, \tag{3.90}$$

with $\beta_1 > 0$ independent of the mesh width. Note that by Lemma 3.87 it holds $(q^h - \overline{q}^h) \in L_0^2(\Omega)$. First, it becomes clear from this representation that (3.90) cannot hold for pairs of finite element spaces for which a function $q^h \in Q^h$ can be found that satisfies (3.76), which is a sufficient condition for V^h/Q^h to violate the discrete inf-sup condition. In (3.76), the left-hand side of (3.90) is zero whereas the right-hand side of (3.90) does not vanish. On the other hand, (3.90) is a weaker condition than the discrete inf-sup condition. Consider for simplicity a quasi-uniform mesh. Then, by properties of the $L^2(\Omega)$ projection (C.28) and the inverse inequality (C.35), one obtains

$$\begin{aligned}\left\|q^h-\overline{q}^h\right\|_Q &= \left(\sum_{K\in\mathcal{T}^h}\left\|q^h-\overline{q}^h\right\|_{L^2(K)}^2\right)^{1/2}\\ &\le C\left(\sum_{K\in\mathcal{T}^h} h_K^2\left\|\nabla q^h\right\|_{L^2(K)}^2\right)^{1/2}\\ &\le C\left(\sum_{K\in\mathcal{T}^h}\left\|q^h\right\|_{L^2(K)}^2\right)^{1/2} = C\left\|q^h\right\|_Q.\end{aligned}$$

Hence, if (3.51) holds, or equivalently (3.52), then (3.90) follows. □

Theorem 3.89 (Sufficient Conditions for the Discrete Inf-Sup Condition Based on an Inf-Sup Condition for Piecewise Constant Finite Element Pressure) *Assume that a discrete inf-sup conditions holds for V^h/P_0 or V^h/Q_0*

$$\sup_{\boldsymbol{v}^h\in V^h\setminus\{\boldsymbol{0}\}} \frac{b^h(\boldsymbol{v}^h,\overline{q}^h)}{\|\boldsymbol{v}^h\|_{V^h}} \ge \beta_0 \left\|\overline{q}^h\right\|_Q \quad \forall\, \overline{q}^h\in P_0(or\ Q_0), \tag{3.91}$$

with $\beta_0 > 0$ *independent of h. Suppose that in addition an inf-sup condition of form* (3.90) *is valid with* β_1 *independent of h. Then, the discrete inf-sup condition* (3.51) *or* (3.52) *holds with*

$$\beta_{\text{is}}^h = \frac{\beta_0\beta_1}{\tilde{C} + \beta_0 + \beta_1} > 0,$$

with $\tilde{C} = 1$ *for conforming velocity spaces and* $\tilde{C} = \sqrt{d}$ *for non-conforming velocity spaces.*

Proof Since the supremum of a sum is lower or equal than the sum of the suprema, one obtains

$$\sup A = \sup (A + B - B) \le \sup (A + B) + \sup B,$$

such that

$$\sup (A + B) \ge \sup A - \sup B. \tag{3.92}$$

Let $q^h \in Q^h$ be arbitrary. Applying (3.92), (3.91), and (3.41) for conforming velocity spaces or (3.40) for non-conforming velocity spaces, respectively, yields

$$\begin{aligned}
\sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b^h(\boldsymbol{v}^h, q^h)}{\|\boldsymbol{v}^h\|_{V^h}} &= \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \left(\frac{b^h(\boldsymbol{v}^h, q^h - \overline{q}^h)}{\|\boldsymbol{v}^h\|_{V^h}} + \frac{b^h(\boldsymbol{v}^h, \overline{q}^h)}{\|\boldsymbol{v}^h\|_{V^h}} \right) \\
&\ge \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b^h(\boldsymbol{v}^h, \overline{q}^h)}{\|\boldsymbol{v}^h\|_{V^h}} - \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b^h(\boldsymbol{v}^h, q^h - \overline{q}^h)}{\|\boldsymbol{v}^h\|_{V^h}} \\
&\ge \beta_0 \left\|\overline{q}^h\right\|_Q - \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{\left\|\nabla \cdot \boldsymbol{v}^h\right\|_{L^2(\Omega)} \left\|q^h - \overline{q}^h\right\|_Q}{\|\boldsymbol{v}^h\|_{V^h}} \\
&\ge \beta_0 \left\|\overline{q}^h\right\|_Q - \tilde{C} \left\|q^h - \overline{q}^h\right\|_Q.
\end{aligned} \tag{3.93}$$

For non-conforming velocity finite element spaces, $\left\|\nabla \cdot \boldsymbol{v}^h\right\|_{L^2(\Omega)}$ has to be replaced by the sum over the mesh cells and (3.40) has to be applied for each mesh cell.

Now, using (3.90) and (3.93) gives

$$\begin{aligned}
\sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b^h(\boldsymbol{v}^h, q^h)}{\|\boldsymbol{v}^h\|_{V^h}} &\ge \beta_1 \left\|q^h - \overline{q}^h\right\|_Q \\
&\ge \beta_1 \left(\frac{\beta_0}{\tilde{C}} \left\|\overline{q}^h\right\|_Q - \frac{1}{\tilde{C}} \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b^h(\boldsymbol{v}^h, q^h)}{\|\boldsymbol{v}^h\|_{V^h}} \right).
\end{aligned}$$

Rearranging terms leads to

$$\sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b^h(\boldsymbol{v}^h, q^h)}{\|\boldsymbol{v}^h\|_{V^h}} \geq \frac{\beta_0 \beta_1}{\tilde{C} + \beta_1} \left\|\overline{q}^h\right\|_Q . \tag{3.94}$$

Using the triangle inequality in (3.90) and inserting (3.94) yields

$$\begin{aligned} \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b^h(\boldsymbol{v}^h, q^h)}{\|\boldsymbol{v}^h\|_{V^h}} &\geq \beta_1 \left\|q^h\right\|_Q - \beta_1 \left\|\overline{q}^h\right\|_Q \\ &\geq \beta_1 \left\|q^h\right\|_Q - \frac{\tilde{C} + \beta_1}{\beta_0} \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b^h(\boldsymbol{v}^h, q^h)}{\|\boldsymbol{v}^h\|_{V^h}} . \end{aligned}$$

Transferring the last term to the left-hand side leads, since q^h was chosen to be arbitrary, to the discrete inf-sup condition (3.52) with the constant β_{is}^h given in the statement of the theorem. ■

Remark 3.90 (To Condition (3.90)*)*

- A so-called patch test for checking condition (3.90) is proposed in Brezzi and Bathe (1990).
- In Brezzi and Bathe (1990), another case was presented where (3.90) leads to the fulfillment of the discrete inf-sup condition, see Theorem 3.96.

□

3.5.3 An Approach for Conforming Velocity Spaces and Continuous Pressure Spaces

Remark 3.91 (Contents) This section presents two criteria for checking the discrete inf-sup condition (3.51) or (3.52) in the case that the pressure space consists of continuous finite element functions. The first criterion was proved in Verfürth (1984) and the second one in Brezzi and Bathe (1990). It turns out that both criteria are more or less equivalent.

The techniques of this section for deriving the discrete inf-sup condition apply an integration by parts of a term of the form $b^h(\boldsymbol{v}, q^h)$ for a continuous function $\boldsymbol{v}$. This step gives in general contributions of jumps of the discrete pressure across faces of mesh cells. The smoothness assumption $Q^h \subset H^1(\Omega)$ guarantees that these jumps vanish.

It will be assumed that $V^h \subset V$ such that $b^h(\cdot,\cdot) = b(\cdot,\cdot)$ and $\|\cdot\|_{V^h} = \|\cdot\|_V$. □

Lemma 3.92 (Lower Bound for the Supremum) *Given a quasi-uniform family of triangulations, let* $V^h \subset V$*, and* $Q^h \subset H^1(\Omega)$*. Assume that there exists a linear operator* $P_1^h : V \to V^h$ *such that*

$$\left\| \boldsymbol{v} - P_1^h \boldsymbol{v} \right\|_{H^k(\Omega)} \le C \left(\sum_{K \in \mathcal{T}^h} h_K^{2-2k} \left\| \nabla \boldsymbol{v} \right\|_{L^2(K)}^2 \right)^{1/2} \quad \forall\ \boldsymbol{v} \in V,\ k = 0, 1, \tag{3.95}$$

with a positive constant C *that does not depend on the mesh width. Then, there exist constants* $C_1, C_2 > 0$ *such that for every* $q^h \in Q^h$

$$\sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}^h, q^h)}{\|\boldsymbol{v}^h\|_V} \ge C_1 \left\| q^h \right\|_Q - C_2 \left(\sum_{K \in \mathcal{T}^h} h_K^2 \left\| \nabla q^h \right\|_{L^2(K)}^2 \right)^{1/2}. \tag{3.96}$$

Proof Let $q^h \in Q^h \subset Q$ be arbitrary. From Corollary 3.44 it follows that there is a $\boldsymbol{v} \in V$ such that $\nabla \cdot \boldsymbol{v} = q^h$ and $\|\boldsymbol{v}\|_V \le C_3 \left\| q^h \right\|_Q$ with C_3 independent of $\boldsymbol{v}$ and q^h. Then, for this $\boldsymbol{v}$, one obtains with the triangle inequality in the denominator, the application of (3.95) for $k = 1$ in the denominator, and the triangle inequality in the numerator

$$\begin{aligned}
\sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}^h, q^h)}{\|\boldsymbol{v}^h\|_V} &\ge \frac{\left| b(P_1^h \boldsymbol{v}, q^h) \right|}{\left\| P_1^h \boldsymbol{v} \right\|_V} \ge \frac{\left| b(P_1^h \boldsymbol{v}, q^h) \right|}{\|\boldsymbol{v}\|_V + \left\| \boldsymbol{v} - P_1^h \boldsymbol{v} \right\|_V} \ge \frac{1}{1 + C} \frac{\left| b(P_1^h \boldsymbol{v}, q^h) \right|}{\|\boldsymbol{v}\|_V} \\
&\ge \frac{1}{1 + C} \left(\frac{\left| b(\boldsymbol{v}, q^h) \right|}{\|\boldsymbol{v}\|_V} - \frac{\left| b(\boldsymbol{v} - P_1^h \boldsymbol{v}, q^h) \right|}{\|\boldsymbol{v}\|_V} \right).
\end{aligned} \tag{3.97}$$

For the first term, one gets

$$\frac{\left| b(\boldsymbol{v}, q^h) \right|}{\|\boldsymbol{v}\|_V} = \frac{\left\| q^h \right\|_Q^2}{\|\boldsymbol{v}\|_V} \ge \frac{1}{C_3} \left\| q^h \right\|_Q.$$

For the second term, integration by parts is applied, leading to

$$\begin{aligned}
\left(\nabla \cdot \left(\boldsymbol{v} - P_1^h \boldsymbol{v} \right), q^h \right) &= \sum_{K \in \mathcal{T}^h} \left(\nabla \cdot \left(\boldsymbol{v} - P_1^h \boldsymbol{v} \right), q^h \right)_K \\
&= \sum_{K \in \mathcal{T}^h} \left[- \left(\nabla q^h, \boldsymbol{v} - P_1^h \boldsymbol{v} \right)_K + \int_{\partial K} \left(q^h \left(\boldsymbol{v} - P_1^h \boldsymbol{v} \right) \boldsymbol{n}_{\partial K} \right)(\boldsymbol{s})\, d\boldsymbol{s} \right].
\end{aligned}$$

The conditions on the regularity of the discrete velocity and pressure imply that terms on the faces cancel out. Applying the Cauchy–Schwarz inequality (A.10), the Cauchy–Schwarz inequality for sums (A.2), (3.95) for $k = 0$, and the quasi-uniformity of the family of meshes yields

$$\begin{aligned}
&\frac{\left|b(\boldsymbol{v} - P_1^h \boldsymbol{v}, q^h)\right|}{\|\boldsymbol{v}\|_V} \\
&= \frac{\left|-(\nabla q^h, \boldsymbol{v} - P_1^h \boldsymbol{v})\right|}{\|\boldsymbol{v}\|_V} \\
&\le \frac{\sum_{K\in\mathcal{T}^h} \left(h_K \left\|\nabla q^h\right\|_{L^2(K)} h_K^{-1} \left\|\boldsymbol{v} - P_1^h \boldsymbol{v}\right\|_{L^2(K)}\right)}{\|\boldsymbol{v}\|_V} \\
&\le \frac{\left(\sum_{K\in\mathcal{T}^h} h_K^2 \left\|\nabla q^h\right\|_{L^2(K)}^2\right)^{1/2} \left(\sum_{K\in\mathcal{T}^h} h_K^{-2} \left\|\boldsymbol{v} - P_1^h \boldsymbol{v}\right\|_{L^2(K)}^2\right)^{1/2}}{\|\boldsymbol{v}\|_V} \\
&\le \left(\min_{K\in\mathcal{T}^h} h_K\right)^{-1} \frac{\left(\sum_{K\in\mathcal{T}^h} h_K^2 \left\|\nabla q^h\right\|_{L^2(K)}^2\right)^{1/2} \left\|\boldsymbol{v} - P_1^h \boldsymbol{v}\right\|_{L^2(\Omega)}}{\|\boldsymbol{v}\|_V} \\
&\le C \left(\min_{K\in\mathcal{T}^h} h_K\right)^{-1} \frac{\left(\sum_{K\in\mathcal{T}^h} h_K^2 \left\|\nabla q^h\right\|_{L^2(K)}^2\right)^{1/2} \left(\sum_{K\in\mathcal{T}^h} h_K^2 \left\|\nabla \boldsymbol{v}\right\|_{L^2(K)}^2\right)^{1/2}}{\|\boldsymbol{v}\|_V} \\
&\le C \left(\min_{K\in\mathcal{T}^h} h_K\right)^{-1} \left(\max_{K\in\mathcal{T}^h} h_K\right) \frac{\left(\sum_{K\in\mathcal{T}^h} h_K^2 \left\|\nabla q^h\right\|_{L^2(K)}^2\right)^{1/2} \left\|\nabla \boldsymbol{v}\right\|_{L^2(\Omega)}}{\|\boldsymbol{v}\|_V} \\
&\le C \left(\sum_{K\in\mathcal{T}^h} h_K^2 \left\|\nabla q^h\right\|_{L^2(K)}^2\right)^{1/2}.
\end{aligned}$$

Inserting these estimates in (3.97) gives (3.96) with $C_1 = 1/(C_3(1 + C))$ and $C_2 = C/(1 + C)$. ■

Remark 3.93 (A Possible Choice of P_1^h) It was already noted in Verfürth (1984) that the Clément operator P_{Cle}^h (C.17), with its modification that preserves homogeneous Dirichlet boundary conditions from Remark C.22, can be used as P_1^h. □

Theorem 3.94 (Criterion for Checking the Discrete Inf-Sup Condition (3.51) for Continuous Pressure Spaces from Verfürth (1984)) *Suppose that the assumptions of Lemma 3.92 hold. Assume in addition that an inf-sup condition of the form*

$$\sup_{\boldsymbol{v}^h\in V^h\setminus\{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}^h, q^h)}{\|\boldsymbol{v}^h\|_V} \ge \beta_2 \left(\sum_{K\in\mathcal{T}^h} h_K^2 \left\|\nabla q^h\right\|_{L^2(K)}^2\right)^{1/2} \quad \forall\, q^h \in Q^h, \tag{3.98}$$

is valid. Then, the discrete inf-sup condition (3.51), or equivalently (3.52), holds.

Proof Adding β_2 times (3.96) and C_2 times (3.98) gives

$$(\beta_2 + C_2) \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}^h, q^h)}{\|\boldsymbol{v}^h\|_V} \geq C_1 \beta_2 \left\| q^h \right\|_Q ,$$

such that (3.52) is satisfied with $\beta_{\text{is}}^h = C_1\beta_2/(\beta_2 + C_2)$. ■

Corollary 3.95 (Criterion for Satisfaction of (3.98)) *Consider a quasi-uniform family of triangulations. If for every $q^h \in Q^h$ there is a $\boldsymbol{v}^h \in V^h$ and positive constants C_1, C_2 such that*

$$b(\boldsymbol{v}^h, q^h) \geq C_1 \left\| \nabla q^h \right\|_{L^2(\Omega)}^2 \quad \text{and} \quad \left\| \boldsymbol{v}^h \right\|_{L^2(\Omega)} \leq C_2 \left\| \nabla q^h \right\|_{L^2(\Omega)} , \tag{3.99}$$

then (3.98) is satisfied.

Proof One obtains with the inverse inequality (C.37) and the assumptions (3.99)

$$\begin{aligned} \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}^h, q^h)}{\|\boldsymbol{v}^h\|_V} &\geq C \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}^h, q^h)}{h^{-1} \|\boldsymbol{v}^h\|_{L^2(\Omega)}} \\ &\geq Ch \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{C_1 \left\| \nabla q^h \right\|_{L^2(\Omega)}^2}{C_2 \|\nabla q^h\|_{L^2(\Omega)}} = Ch \left\| \nabla q^h \right\|_{L^2(\Omega)} , \end{aligned}$$

which is equivalent to (3.98) for families of quasi-uniform triangulations. ■

Theorem 3.96 (Criterion for Checking the Discrete Inf-Sup Condition (3.51) for Continuous Pressure Spaces from Brezzi and Bathe (1990)) *Let $V^h \subset V$, $Q^h \subset H^1(\Omega)$, and consider a quasi-uniform family of triangulations. In addition, suppose that for each $\boldsymbol{v} \in V$ an interpolant $\boldsymbol{v}_I^h \in V^h$ exists such that*

$$\left\| \boldsymbol{v} - \boldsymbol{v}_I^h \right\|_{L^2(\Omega)} \leq Ch \left\| \boldsymbol{v} \right\|_V , \tag{3.100}$$

$$\left\| \boldsymbol{v}_I^h \right\|_V \leq C \left\| \boldsymbol{v} \right\|_V , \tag{3.101}$$

with constants independent of the triangulation. If in addition (3.90) holds, then V^h/Q^h satisfies the discrete inf-sup condition (3.51) or equivalently (3.52).

Proof Under the conditions on V^h and Q^h, integration by parts gives for all $\boldsymbol{v} \in V$

$$\begin{aligned} b(\boldsymbol{v}_I^h - \boldsymbol{v}, q^h) &= \sum_{K \in \mathcal{T}^h} \int_K \nabla \cdot \left(\boldsymbol{v}_I^h - \boldsymbol{v}\right) q^h \, d\boldsymbol{x} \\ &= \sum_{K \in \mathcal{T}^h} \left(\int_{\partial K} q^h \left(\boldsymbol{v}_I^h - \boldsymbol{v}\right) \cdot \boldsymbol{n}_{\partial K} \, ds - \int_K \left(\boldsymbol{v}_I^h - \boldsymbol{v}\right) \cdot \nabla q^h \, d\boldsymbol{x} \right) \\ &= -(\nabla q^h, \boldsymbol{v}_I^h - \boldsymbol{v}). \end{aligned} \tag{3.102}$$

Let $q^h \in Q^h \subset Q$ be arbitrary. Since the pair V/Q satisfies the continuous inf-sup condition (3.48), there is a $\tilde{\boldsymbol{v}} \in V$ such that

$$\frac{b(\tilde{\boldsymbol{v}}, q^h)}{\|\tilde{\boldsymbol{v}}\|_V} \geq (\beta_{\mathrm{is}} - \varepsilon) \left\|q^h\right\|_Q \tag{3.103}$$

with arbitrary small $\varepsilon \in (0, \beta_{\mathrm{is}}/2)$. Denote $\tilde{\beta}_{\mathrm{is}} = \beta_{\mathrm{is}} - \varepsilon > 0$. Taking a special element in the supremum, noting that the supremum is larger or equal also in the case that the negative of this element is taken, applying (3.101), the triangle inequality, (3.103), (3.102), the Cauchy–Schwarz inequality (A.10), and (3.100), it follows that

$$\begin{aligned}
\sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}^h, q^h)}{\|\boldsymbol{v}^h\|_V} &\geq \frac{\left|b(\tilde{\boldsymbol{v}}_I^h, q^h)\right|}{\left\|\tilde{\boldsymbol{v}}_I^h\right\|_V} \geq C \frac{\left|b(\tilde{\boldsymbol{v}}_I^h, q^h)\right|}{\|\tilde{\boldsymbol{v}}\|_V} \\
&\geq C\left(-\frac{\left|b(\tilde{\boldsymbol{v}}_I^h - \tilde{\boldsymbol{v}}, q^h)\right|}{\|\tilde{\boldsymbol{v}}\|_V} + \frac{\left|b(\tilde{\boldsymbol{v}}, q^h)\right|}{\|\tilde{\boldsymbol{v}}\|_V}\right) \\
&\geq C\left(-\frac{\left|b(\tilde{\boldsymbol{v}}_I^h - \tilde{\boldsymbol{v}}, q^h)\right|}{\|\tilde{\boldsymbol{v}}\|_V} + \tilde{\beta}_{\mathrm{is}} \left\|q^h\right\|_Q\right) \\
&= C\left(-\frac{\left|-(\nabla q^h, \tilde{\boldsymbol{v}}_I^h - \tilde{\boldsymbol{v}})\right|}{\|\tilde{\boldsymbol{v}}\|_V} + \tilde{\beta}_{\mathrm{is}} \left\|q^h\right\|_Q\right) \\
&\geq C\left(\tilde{\beta}_{\mathrm{is}} \left\|q^h\right\|_Q - \frac{\left\|\tilde{\boldsymbol{v}}_I^h - \tilde{\boldsymbol{v}}\right\|_{L^2(\Omega)} \left\|\nabla q^h\right\|_{L^2(\Omega)}}{\|\tilde{\boldsymbol{v}}\|_V}\right) \\
&\geq C\left(\tilde{\beta}_{\mathrm{is}} \left\|q^h\right\|_Q - h \left\|\nabla q^h\right\|_{L^2(\Omega)}\right).
\end{aligned} \tag{3.104}$$

Let $\overline{q}^h$ be defined by (3.89). Then, using $\nabla \left(\overline{q}^h\big|_K\right) = 0$ for all mesh cells, the inverse inequality (C.35), and the quasi-uniformity of the family of triangulations gives

$$\begin{aligned}
\left\|\nabla q^h\right\|_{L^2(\Omega)} &= \left(\sum_{K \in \mathcal{T}^h} \left\|\nabla(q^h - \overline{q}^h)\right\|_{L^2(K)}^2\right)^{1/2} \\
&\leq C\left(\sum_{K \in \mathcal{T}^h} h_K^{-2} \left\|q^h - \overline{q}^h\right\|_{L^2(K)}^2\right)^{1/2} \leq C h^{-1} \left\|q^h - \overline{q}^h\right\|_Q.
\end{aligned}$$

Inserting this estimate in (3.104) and using (3.90) yields

$$\sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}^h, q^h)}{\|\boldsymbol{v}^h\|_V} \geq C\tilde{\beta}_{\mathrm{is}} \left\|q^h\right\|_Q - \frac{C}{\beta_1} \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}^h, q^h)}{\|\boldsymbol{v}^h\|_V}.$$

Transferring the last term to the left-hand side and dividing by the arising factor on the left-hand side, finishes the proof. ■

Remark 3.97 (Equivalence of Assumptions (3.90) *and* (3.98)*)* Analogously as in the proof of Theorem 3.96, one can show, using the inverse inequality (C.35), that

$$\left(\sum_{K\in\mathcal{T}^h} h_K^2 \left\|\nabla q^h\right\|_{L^2(K)}^2\right)^{1/2} = \left(\sum_{K\in\mathcal{T}^h} h_K^2 \left\|\nabla (q^h - \overline{q}^h)\right\|_{L^2(K)}^2\right)^{1/2} \le C_{\text{inv}} \left\|q^h - \overline{q}^h\right\|_Q .$$

For the $L^2(\Omega)$ projection into the space of piecewise constant finite element functions, it follows from (C.28) that

$$\left\|q^h - \overline{q}^h\right\|_Q = \left(\sum_{K\in\mathcal{T}^h} \left\|q^h - \overline{q}^h\right\|_{L^2(K)}^2\right)^{1/2} \le C\left(\sum_{K\in\mathcal{T}^h} h_K^2 \left\|\nabla q^h\right\|_{L^2(K)}^2\right)^{1/2} .$$

Hence, the assumptions (3.90) and (3.98) are equivalent. In particular, the discussion of (3.90) from Remark 3.88 applies also to (3.98). □

Remark 3.98 (Theorem 3.94 Implies Theorem 3.96) If the assumptions of Theorem 3.94 are fulfilled, then the assumptions of Theorem 3.96 are satisfied, too. The equivalence of (3.90) and (3.98) was already discussed in Remark 3.97. From the assumptions (3.95) on the operator P_1^h, it follows for quasi-uniform families of triangulations that

$$\left\|\boldsymbol{v} - P_1^h \boldsymbol{v}\right\|_{L^2(\Omega)} \le Ch \left\|\nabla \boldsymbol{v}\right\|_{L^2(\Omega)} = Ch \left\|\boldsymbol{v}\right\|_V ,$$

and

$$\begin{aligned}\left\|P_1^h \boldsymbol{v}\right\|_V &= \left\|\nabla P_1^h \boldsymbol{v}\right\|_{L^2(\Omega)} \le \left\|\nabla (\boldsymbol{v} - P_1^h \boldsymbol{v})\right\|_{L^2(\Omega)} + \left\|\nabla \boldsymbol{v}\right\|_{L^2(\Omega)} \\ &\le (C+1) \left\|\nabla \boldsymbol{v}\right\|_V = C \left\|\nabla \boldsymbol{v}\right\|_V .\end{aligned}$$

Thus, setting $\boldsymbol{v}_I^h = P_1^h \boldsymbol{v}$, the assumptions (3.100) and (3.101) in Theorem 3.96 are also fulfilled. □

3.5.4 Macroelement Techniques

Remark 3.99 (Goal) The goal of macroelement techniques consists in reducing the proof of the discrete inf-sup condition to the proof of a local inf-sup condition. Approaches of this kind were proposed in Boland and Nicolaides (1983) and Stenberg (1984, 1987, 1990). □

Remark 3.100 (The Approach of Boland and Nicolaides (1983)) This approach relies on a decomposition of Ω into open subdomains Ω_r, $r = 1, \ldots, R$, such that

$$\overline{\Omega} = \bigcup_{r=1}^{R} \overline{\Omega_r}, \quad \Omega_r \cap \Omega_s = \emptyset \quad \text{for } r \neq s.$$

Using the global finite element spaces V^h and Q^h, one defines local spaces on Ω_r

$$V^h(\Omega_r) = \left\{ \boldsymbol{v}^h\big|_{\Omega_r} \ : \ \boldsymbol{v}^h \in V^h, \boldsymbol{v}^h = \boldsymbol{0} \text{ in } \Omega \setminus \Omega_r \right\},$$
$$Q^h(\Omega_r) = \left\{ q^h\big|_{\Omega_r} \ : \ q^h \in Q^h \right\} \cap L_0^2(\Omega_r).$$

Then, also the global pressure finite element space consisting of piecewise constant functions is needed

$$Q_0^h = \left\{ q^h \in L_0^2(\Omega) \ : \ q^h\big|_{\Omega_r} \in P_0(\Omega_r), r = 1, \ldots, R \right\}.$$

Now, a local discrete inf-sup condition can be formulated: it exists $\beta_{\text{is,loc}}^h$ independent of h such that

$$\inf_{q^h \in Q^h(\Omega_r)\setminus\{0\}} \ \sup_{\boldsymbol{v}^h \in V^h(\Omega_r)\setminus\{\boldsymbol{0}\}} \frac{\left(\nabla \cdot \boldsymbol{v}^h, q^h\right)_{\Omega_r}}{\|\boldsymbol{v}^h\|_{V^h(\Omega_r)} \|q^h\|_{Q^h(\Omega_r)}} \geq \beta_{\text{is,loc}}^h > 0, \tag{3.105}$$

for all $r = 1, \ldots, R$. □

Theorem 3.101 (Criterion from Boland and Nicolaides (1983)) *Let the local inf-sup condition (3.105) be satisfied with a constant $\beta_{\text{is,loc}}^h > 0$ independent of r and h. Assume that there is a subspace $\overline{V}^h \subset V^h$ such that $\overline{V}^h / Q_0^h$ is inf-sup stable with a constant $\overline{\beta}_{\text{is}}^h > 0$ independent of h. Then, there exists a constant β_{is}^h such that the discrete inf-sup condition (3.51) is satisfied for V^h/Q^h independently of h.*

Proof The proof can be found in Boland and Nicolaides (1983) or Girault and Raviart (1986, Chap. II, Theorem 1.12). ■

Definition 3.102 (Macroelement) Two mesh cells in $\mathbb{R}^d$ are said to be neighbors, if there intersection is a common $d-1$ face. The union of one or more neighboring mesh cells is called a macroelement M. A macroelement M is said to be equivalent to a reference macroelement $\hat{M}$ if there is a map $F_M : \hat{M} \to M$ satisfying the following conditions:

i) F_M is continuous and one-to-one,
ii) $F_M(\hat{M}) = M$,
iii) if $\hat{M} = \cup_{j=1}^m \hat{K}_j$, where $\hat{K}_j, j = 1, \ldots, m$, are the mesh cells contained in $\hat{M}$, then $K_j = F_M(\hat{K}_j), j = 1, \ldots, m$, are the mesh cells defining M,
iv) $F_M|_{\hat{K}_j} = F_{K_j} \circ F_{\hat{K}_j}^{-1}, j = 1, \ldots, m$, where $F_{\hat{K}_j}$ and F_{K_j} are the affine or d-linear maps from the reference mesh cell onto $\hat{K}_j$ and K_j, respectively.

The family of macroelements which are equivalent to $\hat{M}$ will be denoted by $\mathcal{F}_{\hat{M}}$. □

Remark 3.103 (On Macroelements) The meaning of the first two conditions is clear. The third condition states that the image of the restriction of F_M to any mesh cell in $\hat{M}$ is a mesh cell in M. Condition iv) defines the map F_M with the help of the standard maps to reference cells. The restriction of the map F_M to $\hat{K}_j$ is the composition of the reference map from $\hat{K}_j$ to the reference cell $\hat{K}$ and the reference map from $\hat{K}$ to K_j. □

Remark 3.104 (Function Spaces on Macroelements) On the macroelement M, a finite element space is considered whose functions vanish at the boundary of M

$$V_{0,M}^h = \left\{ \boldsymbol{v}^h \in H_0^1(M) \ : \ \boldsymbol{v}^h\big|_K \in R_k(K) \ \forall \ K \subset M \right\}, \tag{3.106}$$

where $R_k(K)$ is $P_k(K)$ if K is a simplex and $Q_k(K)$ if K is a quadrilateral or a parallelepiped. With respect to the pressure, continuous and discontinuous approximations can be considered

$$Q_M^h = \left\{ q^h \in L^2(M) \cap C(\overline{M}) \ : \ q^h\big|_K \in R_l(K) \ \forall \ K \subset M \right\}, \text{ or}$$
$$Q_M^h = \left\{ q^h \in L^2(M) \ : \ q^h\big|_K \in R_l(K) \ \forall \ K \subset M \right\}, \text{ or}$$
$$Q_M^h = \left\{ q^h \in L^2(M) \ : \ q^h\big|_K \in P_l(K) \ \forall \ K \subset M \right\}.$$

Note that in the last spaces the functions from $P_l(K)$ might be defined on quadrilaterals or hexahedra. From these spaces, the functions with vanishing integral mean value are of interest

$$Q_{0,M}^h = Q_M^h \cap L_0^2(M).$$

Note that the finite element spaces defined on M are conforming such that the restriction of the bilinear form $b^h(\cdot, \cdot)$ to M is the same as the restriction of $b(\cdot, \cdot)$ to M and it will be denoted by $b_M(\cdot, \cdot)$.

Finally, a space which corresponds to the kernel of the discrete gradient operator on the macroelement is introduced

$$Q_{\text{grad},M}^h = \left\{ q^h \in Q_M^h \ : \ b_M(\nabla \cdot \boldsymbol{v}^h, q^h) = 0 \ \forall \ \boldsymbol{v}^h \in V_{0,M}^h \right\}.$$

It is clear that the constant functions of Q_M^h are contained in $Q_{\mathrm{grad},M}^h$. □

Remark 3.105 (On the Triangulations) The presentation of the analysis will be restricted for simplicity to the two-dimensional case.

A quasi-uniformity of the family of triangulations has to be assumed, see Definition C.9. For quadrilaterals, in addition it is supposed that

$$|\cos(\theta_{i,K})| \le C, \quad C \in (0,1), \tag{3.107}$$

for all mesh cells K with C independent of K, where $\theta_{i,K}$, $i = 1, \ldots, 4$, are the angles of K. □

Lemma 3.106 (Discrete Inf-Sup Condition on Macroelements) *Let $\mathcal{F}_{\hat{M}}$ be a class of macroelements. Suppose that for every $M \in \mathcal{F}_{\hat{M}}$ the space $Q_{\mathrm{grad},M}^h$ is one-dimensional, consisting only of functions that are constant on M. Then, there is a positive constant $\beta_{\hat{M}}^h$, depending only on $\hat{M}$ and the regularity of the family of triangulations, such that*

$$\sup_{\boldsymbol{v}^h \in V_{0,M}^h \setminus \{\boldsymbol{0}\}} \frac{b_M(\boldsymbol{v}^h, q^h)}{\|\nabla \boldsymbol{v}^h\|_{L^2(M)}} \ge \beta_{\hat{M}}^h \left\|q^h\right\|_{L^2(M)} \quad \forall\, q^h \in Q_{0,M}^h. \tag{3.108}$$

Proof The proof consists of two steps. In the first step, an inf-sup condition of form (3.108) is proved for all $M \in \mathcal{F}_{\hat{M}}$ where the constants depend on M. Then, in the second step, it is shown that these constants can be bounded from below by a constant that depends only on $\hat{M}$ and the regularity of the family of triangulations.

Consider a macroelement $M \in \mathcal{F}_{\hat{M}}$ and define

$$\beta_M^h := \inf_{q^h \in Q_{0,M}^h, \|q^h\|_{L^2(M)}=1} \ \sup_{\boldsymbol{v}^h \in V_{0,M}^h, \|\nabla \boldsymbol{v}^h\|_{L^2(M)}=1} b_M(\boldsymbol{v}^h, q^h) \ge 0. \tag{3.109}$$

This is just the inf-sup condition for the macroelement. The normalization is achieved by including the two factors in the denominator of the usual form of the inf-sup condition into the bilinear form.

The set $\left\{\boldsymbol{v}^h \in V_{0,M}^h \text{ with } \left\|\nabla \boldsymbol{v}^h\right\|_{L^2(M)} = 1\right\}$ is bounded and closed, hence it is a compact set since $V_{0,M}^h$ is a finite-dimensional space. Since $b_M(\cdot,\cdot)$ is continuous, it follows from the Weierstrass Theorem that there is a function $\tilde{\boldsymbol{v}}_{q^h}^h$ such that for each q^h the supremum is attained

$$\sup_{\boldsymbol{v}^h \in V_{0,M}^h, \|\nabla \boldsymbol{v}^h\|_{L^2(M)}=1} b_M(\boldsymbol{v}^h, q^h) = b_M(\tilde{\boldsymbol{v}}_{q^h}^h, q^h) \ge 0.$$

For the same reasons as for the velocity, also $\left\{q^h \in Q_{0,M}^h \text{ with } \left\|q^h\right\|_{L^2(M)} = 1\right\}$ is a compact set and there is a $\tilde{q}^h$ such that

$$\inf_{q^h \in Q_{0,M}^h, \|q^h\|_{L^2(M)}=1} b_M(\tilde{\boldsymbol{v}}_{q^h}^h, q^h) = b_M(\tilde{\boldsymbol{v}}_{q^h}^h, \tilde{q}^h) = \beta_M^h \ge 0.$$

Assume that $\beta_M^h = 0$. Then, one gets for $\tilde{q}^h$

$$\sup_{\boldsymbol{v}^h \in V_{0,M}^h \setminus \{\boldsymbol{0}\}} \frac{b_M(\boldsymbol{v}^h, \tilde{q}^h)}{\|\nabla \boldsymbol{v}^h\|_{L^2(M)}} = \sup_{\boldsymbol{v}^h \in V_{0,M}^h, \|\nabla \boldsymbol{v}^h\|_{L^2(M)}=1} b_M(\boldsymbol{v}^h, \tilde{q}^h) = 0,$$

from what follows that

$$b_M(\boldsymbol{v}^h, \tilde{q}^h) = 0 \quad \forall\ \boldsymbol{v}^h \in V_{0,M}^h.$$

Hence, $\tilde{q}^h \in Q_{\text{grad},M}^h$, i.e., $\tilde{q}^h$ is constant on M. The only constant in $Q_{0,M}^h$ is $\tilde{q}^h = 0$. But then $\left\|\tilde{q}^h\right\|_{L^2(M)} = 0 \neq 1$, which is a contradiction to the definition of the set of functions for taking the infimum. Hence $\beta_M^h > 0$.

Now, it has to be shown that the infimum of β_M^h with respect to M can be bounded uniformly away from zero. To this end, let $\hat{\boldsymbol{x}}_i$, $i = 1, \ldots, d_M$, denote the coordinates of the vertices of the mesh cells in $\hat{M}$. Then, every $M \in \mathcal{F}_{\hat{M}}$ is uniquely determined by the coordinates of its vertices $\boldsymbol{x}_i = F_{\hat{M}}(\hat{\boldsymbol{x}}_i)$, $i = 1, \ldots, d_M$. In this way, the inf-sup constant can be written in the form $\beta_M^h = \beta(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{d_M})$. The coordinates of the vertices can be considered as a point $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{d_M})^T \in \mathbb{R}^{2d_M}$ and one can define a function $\beta(\boldsymbol{X})$ which takes for all admissible coordinates of the vertices of $M \in \mathcal{F}_{\hat{M}}$ the value β_M^h.

Now, set $h_M = \max_{K \in M} h_K$, assume $h_M = 1$, and assume that $\boldsymbol{x}_1$ coincides with the origin in $\mathbb{R}^2$. These assumptions do not restrict generality, since the general situation can be mapped with $\boldsymbol{x} \mapsto h_M^{-1}(\boldsymbol{x} - \boldsymbol{x}_1)$ to this case (translation and scaling). It follows from this construction that all points $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{d_M})$ are within a given distance from the origin. Hence, the arguments $\boldsymbol{X}$ form a bounded set. By the assumptions on the admissible mesh cells, one gets that this set is closed, since in the conditions (C.5) and (3.107) the equal sign leads to admissible mesh cells. Consequently, $\boldsymbol{X}$ is closed and altogether, $\boldsymbol{X}$ is a compact set in $\mathbb{R}^{2d_M}$.

The definition of $\beta(\boldsymbol{X})$, see (3.109), shows that the arguments in the bilinear form change continuously if vertices change continuously, since the arguments are polynomials on each mesh cell, the bilinear form is continuous, the supremum is continuous, and the infimum as well. Altogether, $\beta(\boldsymbol{X})$ defines a continuous function on a compact set. From the first part of the proof it is known that $\beta(\boldsymbol{X}) > 0$ for all arguments. The Theorem of Weierstrass states that a continuous function on a compact set takes its infimum, or, equivalently, the infimum is a minimum. Hence, this minimum is positive and gives the searched inf-sup constant $\beta_{\hat{M}}^h$. ■

Definition 3.107 (The Macroelement Condition) The macroelement condition assumes that there is a fixed set of classes $\mathcal{F}_{\hat{M}_i}$, $i = 1, \ldots, n$, $n \geq 1$, such that for each $M \in \mathcal{F}_{\hat{M}_i}$, $i = 1, \ldots, n$, it is

$$Q_{\text{grad},M}^h \cap Q_{0,M}^h = \{0\}. \tag{3.110}$$

□

Remark 3.108 (On the Macroelement Condition) It will be shown below that the macroelement condition is sufficient for proving that a pair of finite element spaces satisfies the discrete inf-sup condition. For a single mesh cell, (3.110) was already used at the end of the proof of Lemma 3.106. □

Lemma 3.109 (For Each Discrete Pressure Exists a Discrete Velocity with Certain Properties) *Suppose the assumptions of Lemma 3.106 and assume the macroelement condition (3.110). Further, it will be assumed that the mesh cells can be group together to macroelements such that one obtains a partition of $\overline{\Omega}$ consisting of disjoint macroelements with the property that each macroelement belongs to some class $\mathcal{F}_{\hat{M}_i}$, $i = 1, \ldots, n$.*

Then, there is a constant $C_1 > 0$ such that for every $q^h \in Q^h$ there is a function $\boldsymbol{v}_1^h \in V^h$ satisfying

$$\left(\nabla \cdot \boldsymbol{v}_1^h, q^h\right) = \left(\nabla \cdot \boldsymbol{v}_1^h, \left(I - P_{L^2,M}^h\right) q^h\right) = \left\|\left(I - P_{L^2,M}^h\right) q^h\right\|_Q^2 \tag{3.111}$$

and

$$\left\|\boldsymbol{v}_1^h\right\|_V \leq C_1 \left\|\left(I - P_{L^2,M}^h\right) q^h\right\|_Q . \tag{3.112}$$

Here, $P_{L^2,M}^h$ is the $L^2(\Omega)$ projection from Q^h onto the space of piecewise (with respect to the macroelements) constant pressures

$$\tilde{Q}_{0,M}^h = \left\{q^h \in L_0^2(\Omega) \ : \ q^h\big|_M \text{ is constant for all } M\right\} .$$

Proof Note that in general $\left(I - P_{L^2,M}^h\right) q^h \notin Q^h$, e.g., if Q^h is a space consisting of continuous functions. But it is $\left(I - P_{L^2,M}^h\right) q^h\Big|_M \in Q_{0,M}^h$ for all $M \in \mathcal{F}_{\hat{M}_i}$, as shall be proved first. It is $\left(I - P_{L^2,M}^h\right) q^h \in Q_M^h$ since $q^h \in Q_M^h$ and $P_{L^2,M}^h q^h \in Q_M^h$, too. Now, one has to show that the integral mean on M vanishes. By the definition of the projection operator, it is

$$\sum_M r^h\big|_M \int_M \left(I - P_{L^2,M}^h\right) q^h(\boldsymbol{x})\, d\boldsymbol{x} = 0 \quad \forall\ r^h \in \tilde{Q}_{0,M}^h,$$

where the sum is taken over all macroelements. Let $\tilde{M}$ be an arbitrary macroelement. Taking

$$r^h\big|_M = \begin{cases} 1 + c & \text{in } \tilde{M}, \\ c & \text{else}, \end{cases}$$

where $c \in \mathbb{R}$ is chosen such that $\int_\Omega r^h(\boldsymbol{x})\,d\boldsymbol{x} = 0$, gives

$$\begin{aligned}
0 &= (1+c)\int_{\tilde{M}} \left(I - P^h_{L^2,M}\right) q^h(\boldsymbol{x})\,d\boldsymbol{x} + c \sum_{M, M\neq\tilde{M}} \int_M \left(I - P^h_{L^2,M}\right) q^h(\boldsymbol{x})\,d\boldsymbol{x} \\
&= \int_{\tilde{M}} \left(I - P^h_{L^2,M}\right) q^h(\boldsymbol{x})\,d\boldsymbol{x} + c\sum_M \int_M \left(I - P^h_{L^2,M}\right) q^h(\boldsymbol{x})\,d\boldsymbol{x} \\
&= \int_{\tilde{M}} \left(I - P^h_{L^2,M}\right) q^h(\boldsymbol{x})\,d\boldsymbol{x} + c\int_\Omega q^h(\boldsymbol{x})\,d\boldsymbol{x} - c\int_\Omega P^h_{L^2,M} q^h(\boldsymbol{x})\,d\boldsymbol{x}.
\end{aligned}$$

Since $q^h \in L^2_0(\Omega)$ and $P^h_{L^2,M} q^h \in L^2_0(\Omega)$, the last two integrals vanish and one gets $\left(I - P^h_{L^2,M}\right) q^h \in L^2_0(M)$ for an arbitrary macroelement M.

From the inf-sup condition (3.108) on M, it follows that for $\left.\left(I - P^h_{L^2,M}\right) q^h\right|_M$ there is a $\boldsymbol{v}^h_M \in V^h_{0,M}$, more precisely in the orthogonal complement of the discretely divergence-free functions of $V^h_{0,M}$, such that $\nabla\cdot\boldsymbol{v}^h_M = \left.\left(I - P^h_{L^2,M}\right) q^h\right|_M$ and

$$\left\|\nabla \boldsymbol{v}^h_M\right\|_{L^2(M)} \le C_{1,M} \left\|\left(I - P^h_{L^2,M}\right) q^h\right\|_{L^2(M)},$$

which is proved analogously to Corollary 3.44. Analogously to Corollary 3.47, see also Lemma 3.12 iii), one finds that $C_{1,M} = 1/\beta^h_{\hat{M}}$ and setting

$$C_1 = \left(\min_{i=1,\ldots,n} \beta^h_{\hat{M}_i}\right)^{-1}$$

leads to a uniform estimate with constant C_1 for all classes $\mathcal{F}_{\hat{M}_i}$. In addition, it is

$$b_M\left(\boldsymbol{v}^h_M, \left(I - P^h_{L^2,M}\right) q^h\right) = \left(\nabla\cdot\boldsymbol{v}^h_M, \left(I - P^h_{L^2,M}\right) q^h\right)_M = \left\|\left(I - P^h_{L^2,M}\right) q^h\right\|^2_{L^2(M)}.$$

Now, one takes $\boldsymbol{v}^h_1 \in V^h$ with $\left.\boldsymbol{v}^h_1\right|_M = \boldsymbol{v}^h_M$. Summation over all macros gives the equation on the right-hand side of (3.111) and the estimate (3.112).

Since $\boldsymbol{v}^h_1 = \boldsymbol{0}$ on the boundary of every macroelement, one gets

$$\left(\nabla\cdot\boldsymbol{v}^h_1, P^h_{L^2,M} q^h\right) = 0 \quad \forall\, q^h \in Q^h,$$

which gives finally the equation on the left-hand side of (3.111). ■

Lemma 3.110 (For Each Discrete Pressure Exists Another Discrete Velocity with Certain Properties) *Let the assumption of Lemma 3.109 be valid and let $k \geq 2$ in (3.106). Then there is a constant $C_2 > 0$ such that for every $q^h \in Q^h$ there is a $\boldsymbol{v}_2^h \in V^h$ such that*

$$\left(\nabla \cdot \boldsymbol{v}_2^h, P_{L^2,M}^h q^h\right) = \left\| P_{L^2,M}^h q^h \right\|_Q^2 \quad \text{and} \quad \left\| \boldsymbol{v}_2^h \right\|_V \leq C_2 \left\| P_{L^2,M}^h q^h \right\|_Q .$$

Proof Let $q^h \in Q^h$ be arbitrary. Since $P_{L^2,M}^h q^h \in Q$, there is a $\boldsymbol{v} \in V$ such that

$$\nabla \cdot \boldsymbol{v} = P_{L^2,M}^h q^h \quad \text{and} \quad \|\boldsymbol{v}\|_V \leq C \left\| P_{L^2,M}^h q^h \right\|_Q , \tag{3.113}$$

see Lemma 3.12 iii) or Corollary 3.47. Now, an operator $I^h : V \to V^h$ is constructed with

$$\begin{aligned} \left(\nabla \cdot I^h \boldsymbol{v}, \overline{q}^h\right) &= \left(\nabla \cdot \boldsymbol{v}, \overline{q}^h\right) \quad \forall\, \overline{q}^h \in \tilde{Q}_{0,M}^h, \\ \left\| I^h \boldsymbol{v} \right\|_V &\leq C \|\boldsymbol{v}\|_V . \end{aligned}$$

Note that these are the properties of a Fortin operator for $\tilde{Q}_{0,M}^h$, compare (3.82). For details of this construction, it is referred to Stenberg (1984). Setting $\boldsymbol{v}_2^h = I^h \boldsymbol{v}$, $\overline{q}^h = P_{L^2,M}^h q^h$, and using the properties (3.113) of $\boldsymbol{v}$ leads to the statement of the lemma. ■

Theorem 3.111 (Macroelement Condition of Stenberg (1984)) *Let the assumptions of Lemmas 3.109 and 3.110 be satisfied, then the discrete inf-sup condition (3.51) or (3.52) holds.*

Proof Choose $q^h \in Q^h$ arbitrarily, let $\boldsymbol{v}_1^h, \boldsymbol{v}_2^h \in V^h$, and let C_1, C_2 be the constants from Lemmas 3.109 and 3.110. Setting

$$\boldsymbol{v}^h = \boldsymbol{v}_1^h + C_3 \boldsymbol{v}_2^h, \quad C_3 = \frac{2}{1 + C_2^2},$$

one obtains with (3.111)

$$\begin{aligned} \left(\nabla \cdot \boldsymbol{v}^h, q^h\right) &= \left(\nabla \cdot \boldsymbol{v}_1^h, q^h\right) + C_3 \left(\nabla \cdot \boldsymbol{v}_2^h, q^h\right) \\ &= \left\| \left(I - P_{L^2,M}^h\right) q^h \right\|_Q^2 + C_3 \left(\nabla \cdot \boldsymbol{v}_2^h, P_{L^2,M}^h q^h\right) \\ &\quad + C_3 \left(\nabla \cdot \boldsymbol{v}_2^h, \left(I - P_{L^2,M}^h\right) q^h\right) . \end{aligned}$$

Next, Lemma 3.110 is applied to the second term on the right-hand side. The last term is estimated with the Cauchy–Schwarz inequality (A.10), (3.41), Lemma 3.110, and Young's inequality (A.5), leading to

$$
\begin{aligned}
&\left(\nabla \cdot \boldsymbol{v}^h, q^h\right) \\
&\ge \left\| \left(I - P^h_{L^2,M}\right) q^h \right\|_Q^2 + C_3 \left\| P^h_{L^2,M} q^h \right\|_Q^2 - C_3 \left\| \nabla \cdot \boldsymbol{v}_2^h \right\|_{L^2(\Omega)} \left\| \left(I - P^h_{L^2,M}\right) q^h \right\|_Q \\
&\ge \left\| \left(I - P^h_{L^2,M}\right) q^h \right\|_Q^2 + C_3 \left\| P^h_{L^2,M} q^h \right\|_Q^2 - C_3 C_2 \left\| P^h_{L^2,M} q^h \right\|_Q \left\| \left(I - P^h_{L^2,M}\right) q^h \right\|_Q \\
&\ge \left\| \left(I - P^h_{L^2,M}\right) q^h \right\|_Q^2 + C_3 \left\| P^h_{L^2,M} q^h \right\|_Q^2 - \frac{C_3}{2} \left\| P^h_{L^2,M} q^h \right\|_Q^2 \\
&\quad - \frac{C_3 C_2^2}{2} \left\| \left(I - P^h_{L^2,M}\right) q^h \right\|_Q^2 \\
&= \left(1 - \frac{C_3 C_2^2}{2}\right) \left\| \left(I - P^h_{L^2,M}\right) q^h \right\|_Q^2 + \frac{C_3}{2} \left\| P^h_{L^2,M} q^h \right\|_Q^2 .
\end{aligned}
$$

Next, the orthogonality of the $L^2(\Omega)$ projection is used,

$$
\left\| \left(I - P^h_{L^2,M}\right) q^h \right\|_Q^2 = \left\| q^h \right\|_Q^2 + \left\| P^h_{L^2,M} q^h \right\|_Q^2 ,
$$

as well as the definition of C_3, giving

$$
\begin{aligned}
\left(\nabla \cdot \boldsymbol{v}^h, q^h\right) &\ge \left(1 - \frac{C_3 C_2^2}{2}\right) \left\| q^h \right\|_Q^2 + \underbrace{\left(1 - \frac{C_3 C_2^2}{2} + \frac{C_3}{2}\right)}_{>0} \left\| P^h_{L^2,M} q^h \right\|_Q^2 \\
&\ge \frac{1}{1 + C_2^2} \left\| q^h \right\|_Q^2 .
\end{aligned}
$$

Applying the triangle inequality, (3.112), Lemma 3.110, and the stability of the $L^2(\Omega)$ projection (C.26) leads to

$$
\begin{aligned}
\left\| \boldsymbol{v}^h \right\|_V &\le \left\| \boldsymbol{v}_1^h \right\|_V + C_3 \left\| \boldsymbol{v}_2^h \right\|_V \\
&\le C_1 \left\| \left(I - P^h_{L^2,M}\right) q^h \right\|_Q + C_3 C_2 \left\| P^h_{L^2,M} q^h \right\|_Q \\
&\le C \left\| q^h \right\|_Q .
\end{aligned}
$$

Combining the last two estimates gives that for all $q^h \in Q^h$ there is a $\boldsymbol{v}^h \in V^h$ with

$$
\frac{\left(\nabla \cdot \boldsymbol{v}^h, q^h\right)}{\left\| \boldsymbol{v}^h \right\|_V} \ge \frac{1}{1 + C_2^2} \frac{\left\| q^h \right\|_Q^2}{\left\| \boldsymbol{v}^h \right\|_V} \ge \frac{1}{C\left(1 + C_2^2\right)} \left\| q^h \right\|_Q .
$$

Taking the supremum with respect to V^h on both sides of this estimate, one obtains that the discrete inf-sup condition (3.52) is satisfied. ■

Remark 3.112 (Relaxing the Condition on the Macroelements) The condition that the macroelements should be disjoint can be relaxed, e.g., see Stenberg (1987, 1990). It is sufficient that there is a fixed number m such that each mesh cell does not belong to more than m macroelements. □

3.6 Inf-Sup Stable Pairs of Finite Element Spaces

Remark 3.113 (Contents) This section present pairs of inf-sup stable finite element spaces. For some of these pairs, the proof of the discrete inf-sup condition will be given in detail. □

3.6.1 The MINI Element

Remark 3.114 (The MINI Element) The MINI element is defined on simplicial grids and it is given by

$$V^h = P_1 \oplus V_{\text{bub}}^h, \quad Q^h = P_1, \tag{3.114}$$

where V_{bub}^h is a space consisting of local bubble functions

$$V_{\text{bub}}^h = \left\{ v_{\text{bub}}^h \ : \ \text{supp}\left(v_{\text{bub}}^h\right) = K,\, v_{\text{bub}}^h\big|_K = \alpha \prod_{i=1}^{d+1} \lambda_i, K \in \mathcal{T}^h, \alpha \in \mathbb{R} \right\},$$

where λ_i are the barycentric coordinates of the simplex K, see Definition B.31. It follows that

$$v_{\text{bub}}^h\big|_K \in P_{d+1}(K) \cap H_0^1(K).$$

This pair of finite element spaces was introduced by Arnold et al. (1984). It is the lowest order conforming inf-sup stable pair of finite element spaces.

The basic idea for the construction of the MINI element consists in starting with standard finite element spaces for velocity and pressure and then enriching the velocity space such that the discrete inf-sup condition (3.51) is satisfied. The fulfillment of the discrete inf-sup condition will be proved with the construction of a Fortin operator, see Lemma 3.78. □

Lemma 3.115 (Properties of Bubble Functions) *Let $K \in \mathcal{T}^h$ be a simplex and let*

$$v_{\mathrm{bub}}^h(\boldsymbol{x}) = \prod_{i=1}^{d+1} \lambda_i(\boldsymbol{x}), \quad \boldsymbol{x} \in K,$$

be a bubble function on K. Then, the following estimates hold

$$\left\| v_{\mathrm{bub}}^h \right\|_{L^2(K)} \le C h_K^d, \tag{3.115}$$

$$\left\| \nabla v_{\mathrm{bub}}^h \right\|_{L^2(K)} \le C h_K^{(d-2)/2}, \tag{3.116}$$

$$\int_K v_{\mathrm{bub}}^h(\boldsymbol{x})\, d\boldsymbol{x} \ge C\, |K|\,, \tag{3.117}$$

where the constants are independent of K.

Proof • *Estimate* (3.115). This estimate follows directly by Hölder's inequality (A.9)

$$\left\| v_{\mathrm{bub}}^h \right\|_{L^2(K)} \le \left\| v_{\mathrm{bub}}^h \right\|_{L^\infty(K)} \|1\|_{L^1(K)} = \left\| v_{\mathrm{bub}}^h \right\|_{L^\infty(K)} |K| = C h_K^d. \tag{3.118}$$

• *Estimate* (3.116). Since K is a simplex, there is an affine transform (B.18) to the reference simplex $\hat{K}$. Applying this transform to the integral gives, see also (B.34),

$$\begin{aligned} \left\| \nabla v_{\mathrm{bub}}^h \right\|_{L^2(K)}^2 &= \int_K \left(\nabla v_{\mathrm{bub}}^h \cdot \nabla v_{\mathrm{bub}}^h \right)(\boldsymbol{x})\, d\boldsymbol{x} \\ &= |\det(B)| \int_{\hat{K}} B^{-T} \nabla_{\boldsymbol{\xi}}\, \hat{v}_{\mathrm{bub}}^h(\boldsymbol{\xi}) \cdot B^{-T} \nabla_{\boldsymbol{\xi}}\, \hat{v}_{\mathrm{bub}}^h(\boldsymbol{\xi})\, d\boldsymbol{\xi}, \end{aligned}$$

where B is the matrix of the affine transform. Using now the triangle inequality, the Cauchy–Schwarz inequality (A.2), and the compatibility of the Euclidean vector norm and the spectral matrix norm yields

$$\begin{aligned} \left\| \nabla v_{\mathrm{bub}}^h \right\|_{L^2(K)}^2 &\le |\det(B)| \int_{\hat{K}} \left| B^{-T} \nabla_{\boldsymbol{\xi}}\, \hat{v}_{\mathrm{bub}}^h(\boldsymbol{\xi}) \cdot B^{-T} \nabla_{\boldsymbol{\xi}}\, \hat{v}_{\mathrm{bub}}^h(\boldsymbol{\xi}) \right| \, d\boldsymbol{\xi} \\ &\le |\det(B)| \int_{\hat{K}} \left\| B^{-T} \nabla_{\boldsymbol{\xi}}\, \hat{v}_{\mathrm{bub}}^h(\boldsymbol{\xi}) \right\|_2^2 \, d\boldsymbol{\xi} \\ &\le |\det(B)| \left\| B^{-T} \right\|_2^2 \int_{\hat{K}} \left\| \hat{v}_{\mathrm{bub}}^h(\boldsymbol{\xi}) \right\|_2^2 \, d\boldsymbol{\xi}. \end{aligned}$$

The last factor does not depend on K and it can be considered to be a constant. Applying finally (C.6) and (C.9) gives

$$\left\| \nabla v_{\text{bub}}^h \right\|_{L^2(K)}^2 \le C h_K^d h_K^{-2} = C h_K^{d-2}.$$

- *Estimate* (3.117). The bubble functions are polynomials of degree $d+1$ in K. Hence, there are quadrature rules with positive weights and nodes in the interior of K such that they can be integrated exactly

$$\int_K v_{\text{bub}}^h(\boldsymbol{x})\,d\boldsymbol{x} = |K| \sum_{i=1}^{N_0} \omega_i v_{\text{bub}}^h(\boldsymbol{x}_i),$$

see Remark 3.116, i.e., $\omega_i > 0$, $v_{\text{bub}}^h(\boldsymbol{x}_i) > 0$, $i = 1, \ldots, N_0$. It follows that

$$\begin{aligned}\left| \int_K v_{\text{bub}}^h(\boldsymbol{x})\,d\boldsymbol{x} \right| &= |K| \left| \sum_{i=1}^{N_0} \omega_i v_{\text{bub}}^h(\boldsymbol{x}_i) \right| = |K| \sum_{i=1}^{N_0} \omega_i v_{\text{bub}}^h(\boldsymbol{x}_i) \\ &\ge |K| \min_{i=1,\ldots,N_0} v_{\text{bub}}^h(\boldsymbol{x}_i) \sum_{i=1}^{N_0} \omega_i = C\,|K|\,.\end{aligned}$$

■

Remark 3.116 (Concerning the Bubble Functions)

- One can show with standard calculus that a bubble function takes its maximum in the barycenter of K. Then, the factor in (3.118) is

$$\left\| v_{\text{bub}}^h \right\|_{L^\infty(K)} = \frac{1}{27} \ \text{ if } \ d = 2, \quad \left\| v_{\text{bub}}^h \right\|_{L^\infty(K)} = \frac{1}{216} \ \text{ if } \ d = 3.$$

- In Cools and Rabinowitz (1993), one can find that there are quadrature rules with positive weights and nodes in the interior of a simplex for polynomials of degree $2 \le k \le 7$ for $d = 2$ and of degree $1 \le k \le 5$ for $d = 3$. For triangles, the so-called Gauss3 quadrature rule can be used, see Stroud (1971), which is exact for polynomials of order 5 and possesses seven nodes.

□

Remark 3.117 (Generalization of the Fortin Criterion (3.82)) The first part of the Fortin criterion (3.82) can be written in the form

$$-\int_\Omega \nabla \cdot \left(\boldsymbol{v} - P_{\text{For}}^h \boldsymbol{v} \right) q^h\,d\boldsymbol{x} = 0 \quad \forall\ \boldsymbol{v} \in V,\ \forall\ q^h \in Q^h.$$

It follows, for conforming finite element spaces and a continuous finite element pressure space, using integration by parts, that

$$\int_\Omega \left(\boldsymbol{v} - P_{\text{For}}^h \boldsymbol{v} \right) \cdot \nabla q^h\,d\boldsymbol{x} = 0 \quad \forall\ \boldsymbol{v} \in V,\ \forall\ q^h \in Q^h. \tag{3.119}$$

The first step of the construction of the Fortin operator consists in replacing the global criterion (3.119) by a set of local criteria

$$\int_K \left(\boldsymbol{v} - P^h_{\mathrm{For}}\boldsymbol{v}\right) \cdot \nabla q^h \, d\boldsymbol{x} = 0 \quad \forall\, \boldsymbol{v} \in V,\ \forall\, q^h \in Q^h,\ \forall\, K \in \mathcal{T}^h. \tag{3.120}$$

Clearly, (3.120) induces (3.119), but not vice versa. □

Remark 3.118 (Enrichment of the Velocity Space) Let $Q^h(K) = P_k(K)$, then it follows that $\nabla q^h \in P_{k-1}(K)$. It is clear that (3.120) can be satisfied, for fixed Q^h, the easier the larger the space $V^h(K)$ is, since for a larger space $V^h(K)$ there are more possibilities to define $P^h_{\mathrm{For}}\boldsymbol{v}$. The idea of Arnold et al. (1984) was to start for $V^h(K)$ also with polynomials of order k and then to extend this space locally, i.e., with functions whose support is restricted to K, until the velocity space is sufficiently large to satisfy (3.120). □

Remark 3.119 (Local Condition (3.120) *for the MINI Element)* For the MINI element (3.114), condition (3.120) simplifies to

$$\int_K \left(\boldsymbol{v} - P^h_{\mathrm{For}}\boldsymbol{v}\right) \, d\boldsymbol{x} = \boldsymbol{0} \quad \forall\, \boldsymbol{v} \in V,\ \forall\, K \in \mathcal{T}^h, \tag{3.121}$$

since the gradient of the local discrete pressure is a constant. □

Remark 3.120 (Construction of the Fortin Operator) The construction of the Fortin operator is based on the Clément interpolation operator P^h_{Cle} defined (C.17), with the modification to preserve homogeneous Dirichlet boundary conditions, see Remark 3.83. This operator satisfies the interpolation estimate (C.18). Consider a quasi-uniform family of triangulations. Then, the number of mesh cells in the set ω_K from (C.18) is bounded uniformly from above and one gets the global estimates

$$\sum_{K\in\mathcal{T}^h} h_K^{-2} \left\| v - P^h_{\mathrm{Cle}} v \right\|^2_{L^2(K)} \le C \left\| v \right\|^2_{H^1(\Omega)} \quad \forall\, v \in H^1(\Omega), \tag{3.122}$$

$$\sum_{K\in\mathcal{T}^h} \left\| \nabla \left(v - P^h_{\mathrm{Cle}} v\right) \right\|^2_{L^2(K)} \le C \left\| v \right\|^2_{H^1(\Omega)} \quad \forall\, v \in H^1(\Omega). \tag{3.123}$$

From the triangle inequality and (3.123), one gets in particular the stability estimate

$$\begin{aligned} &\sum_{K\in\mathcal{T}^h} \left\| \nabla P^h_{\mathrm{Cle}} v \right\|^2_{L^2(K)} \\ &\le 2 \left(\sum_{K\in\mathcal{T}^h} \left\| \nabla \left(v - P^h_{\mathrm{Cle}} v\right) \right\|^2_{L^2(K)} + \sum_{K\in\mathcal{T}^h} \left\| \nabla v \right\|^2_{L^2(K)} \right) \le C \left\| v \right\|^2_{H^1(\Omega)} . \end{aligned} \tag{3.124}$$

Now, the Fortin operator is defined by

$$P^h_{\mathrm{For}} \boldsymbol{v}(\boldsymbol{x}) = P^h_{\mathrm{Cle}} \boldsymbol{v}(\boldsymbol{x}) + \boldsymbol{\alpha}_K v^h_{\mathrm{bub}}(\boldsymbol{x}), \tag{3.125}$$

with

$$\boldsymbol{\alpha}_K = \frac{\int_K \left(\boldsymbol{v} - P^h_{\mathrm{Cle}} \boldsymbol{v}\right)(\boldsymbol{x})\, d\boldsymbol{x}}{\int_K v^h_{\mathrm{bub}}(\boldsymbol{x})\, d\boldsymbol{x}}. \tag{3.126}$$

This construction is of form (3.87) with P^h_2 just being the integral operator on K equipped with some scaling. □

Theorem 3.121 (The Discrete Inf-Sup Condition for the MINI Element) *Consider a quasi-uniform family of triangulations. Then, the operator* (3.125) *is a Fortin operator. Hence, the MINI element* (3.114) *satisfies the discrete inf-sup condition* (3.51) *or equivalently* (3.52)*.*

Proof One has to verify the conditions stated in (3.82). Instead of the first of these conditions, the more general condition (3.121) will be considered.

- *Condition* (3.121)*.* Inserting (3.125) and (3.126) in (3.121) yields for an arbitrary mesh cell K

$$\begin{aligned}
&\int_K \left(\boldsymbol{v} - P^h_{\mathrm{For}} \boldsymbol{v}\right)(\boldsymbol{x})\, d\boldsymbol{x} \\
&= \int_K \left(\boldsymbol{v} - P^h_{\mathrm{Cle}} \boldsymbol{v} - \boldsymbol{\alpha}_K v^h_{\mathrm{bub}}\right)(\boldsymbol{x})\, d\boldsymbol{x} \\
&= \int_K \left(\boldsymbol{v} - P^h_{\mathrm{Cle}} \boldsymbol{v}\right)(\boldsymbol{x})\, d\boldsymbol{x} - \boldsymbol{\alpha}_K \int_K v^h_{\mathrm{bub}}(\boldsymbol{x})\, d\boldsymbol{x} \\
&= \int_K \left(\boldsymbol{v} - P^h_{\mathrm{Cle}} \boldsymbol{v}\right)(\boldsymbol{x})\, d\boldsymbol{x} - \frac{\int_K \left(\boldsymbol{v} - P^h_{\mathrm{Cle}} \boldsymbol{v}\right)(\boldsymbol{x})\, d\boldsymbol{x}}{\int_K v^h_{\mathrm{bub}}(\boldsymbol{x})\, d\boldsymbol{x}} \int_K v^h_{\mathrm{bub}}(\boldsymbol{x})\, d\boldsymbol{x} = \boldsymbol{0}.
\end{aligned}$$

- *Second condition of* (3.82)*.* The triangle inequality and the homogeneity of a norm, Definition A.6, gives

$$\left\| \nabla P^h_{\mathrm{For}} \boldsymbol{v} \right\|_{L^2(K)} \le \left\| \nabla P^h_{\mathrm{Cle}} \boldsymbol{v} \right\|_{L^2(K)} + \left\| \boldsymbol{\alpha}_K \right\|_2 \left\| \nabla v^h_{\mathrm{bub}} \right\|_{L^2(K)}, \tag{3.127}$$

where $\|\boldsymbol{\alpha}_K\|_2$ is the Euclidean norm of the vector-valued constant $\boldsymbol{\alpha}_K$. One obtains for the second term, using (3.126), (3.116), the Cauchy–Schwarz inequality (A.10), (3.117), and $|K| = Ch^d_K$

$$\begin{aligned}
\left\| \boldsymbol{\alpha}_K \right\|_2 \left\| \nabla v^h_{\mathrm{bub}} \right\|_{L^2(K)} &\le C \frac{\left\| \int_K \left(\boldsymbol{v} - P^h_{\mathrm{Cle}} \boldsymbol{v}\right)(\boldsymbol{x})\, d\boldsymbol{x} \right\|_2}{\left| \int_K v^h_{\mathrm{bub}}(\boldsymbol{x})\, d\boldsymbol{x} \right|} h_K^{(d-2)/2} \\
&\le C \frac{\left\| \boldsymbol{v} - P^h_{\mathrm{Cle}} \boldsymbol{v} \right\|_{L^2(K)} |K|^{1/2}}{|K|} h_K^{(d-2)/2}
\end{aligned}$$

$$\leq C \left\| \boldsymbol{v} - P^h_{\mathrm{Cle}} \boldsymbol{v} \right\|_{L^2(K)} h_K^{d/2-1-d/2}$$
$$= C h_K^{-1} \left\| \boldsymbol{v} - P^h_{\mathrm{Cle}} \boldsymbol{v} \right\|_{L^2(K)} .$$

Inserting this estimate in (3.127), taking the square, using Young's inequality (A.5), and summing over all mesh cells gives

$$\left\| P^h_{\mathrm{For}} \boldsymbol{v} \right\|_V^2 \leq C \left(\sum_{K \in \mathcal{T}^h} \left\| \nabla P^h_{\mathrm{Cle}} \boldsymbol{v} \right\|^2_{L^2(K)} + h_K^{-2} \left\| \boldsymbol{v} - P^h_{\mathrm{Cle}} \boldsymbol{v} \right\|^2_{L^2(K)} \right).$$

Now the proof is finished by inserting (3.124), (3.122), and applying Poincaré's inequality (A.12).

■

Remark 3.122 (To MINI-type Elements)

- Using the MINI element is quite popular.
- The construction of the MINI element can be extended to higher order finite elements, see Arnold et al. (1984). But to the best of our knowledge, the use of these higher order elements is not popular.
- It is mentioned in Boffi et al. (2008, Section 4.6) that almost any pair of finite element spaces can be stabilized by enriching the velocity space with bubble functions.

□

3.6.2 The Family of Taylor–Hood Finite Elements

Remark 3.123 (The Family of Taylor–Hood Finite Element Spaces) The family of Taylor–Hood finite element spaces on triangular and tetrahedral grids is given by P_k/P_{k-1}, $k \geq 2$, and on quadrilateral and hexahedral grids by Q_k/Q_{k-1}, $k \geq 2$. That means, the pressure is approximated by a continuous function. Hence, it is $b^h(\cdot,\cdot) = b(\cdot,\cdot)$ and $\|\cdot\|_{V^h} = \|\cdot\|_V$.

In Hood and Taylor (1974), the use of the $Q_2^{(8)}/Q_1$ pair of finite element spaces was proposed for solving the Navier–Stokes equations on quadrilateral meshes, where the $Q_2^{(8)}$ finite element is the Q_2 finite element without internal degree of freedom, see Fig. 3.6. It was proved in Stenberg (1984) with the macroelement technique that the $Q_2^{(8)}/Q_1$ pair of finite element spaces satisfies the discrete inf-sup condition on meshes consisting of rectangles (such that the domain Ω is just a union of rectangles).

The pairs of Taylor–Hood finite element spaces are among the most popular pairs for discretizing equations modeling incompressible flows, in particular the pairs for $k = 2$. A reason for this popularity is certainly that the implementation of the P_2/P_1

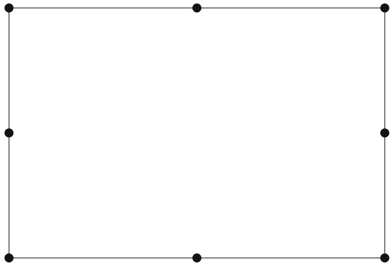

Fig. 3.6 The finite element $Q_2^{(8)}$

and Q_2/Q_1 finite element pairs is comparatively easy compared with other inf-sup stable pairs of finite elements.

Note that it was already proved that the pairs P_1/P_0 and Q_1/Q_0 are not inf-sup stable, see Examples 3.70 and 3.71. □

Remark 3.124 (Historical Remarks Concerning P_k/P_{k-1}, $k \geq 2$) Because of their popularity there were already rather early attempts to prove the discrete inf-sup condition for pairs of Taylor–Hood finite element spaces.

In Bercovier and Pironneau (1979) a discrete inf-sup condition for the pair P_2/P_1 in two dimensions was proved with wrong norms

$$\sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}^h, q^h)}{\|\boldsymbol{v}^h\|_{L^2(\Omega)}} \geq C \left\|\nabla q^h\right\|_{L^2(\Omega)} \quad \forall\, q^h \in Q^h. \tag{3.128}$$

The connections of (3.128) to the discrete inf-sup condition (3.51) are studied in Guzmán et al. (2013).

In proof presented in Bercovier and Pironneau (1979), the square of the gradient of the pressure is estimated by a sum of squares of derivatives on interior edges, see Bercovier and Pironneau (1979, estimate (2.11)). This estimate requires that there are at least two interior edges for each triangle and hence, it was assumed that each triangle possesses a vertex that is not on $\partial\Omega$. In Verfürth (1984), the correct discrete inf-sup condition (3.51) was proved, based on (3.128) and the criterion formulated in Theorem 3.94. The assumptions include still that all triangles have at least one interior vertex. Then, in Stenberg (1990) and Brezzi and Falk (1991) the discrete inf-sup condition was proved for P_3/P_2 in two dimensions. The analysis of Brezzi and Falk (1991) requires still some mild assumption concerning triangles with two edges on the boundary. Finally, a proof for arbitrary order spaces P_k/P_{k-1}, $k \geq 2$, without restrictions with respect to triangles with two edges on $\partial\Omega$ was presented in Boffi (1994). This proof uses also Theorem 3.94. A Fortin operator for the P_2/P_1 pair of spaces in two dimensions was constructed in Mardal et al. (2013).

The inf-sup condition for P_2/P_1 in three dimensions was proved in Stenberg (1987) with the macroelement technique and the analysis was extended to higher order Taylor–Hood pairs in Boffi (1997). It is required in the analysis that each tetrahedron has at least one vertex in the interior of Ω. The presentation in this section follows Boffi et al. (2013, Section 8.8.2) and Boffi (1994). □

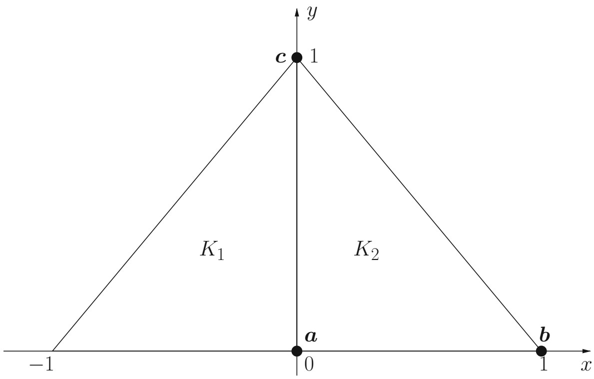

Fig. 3.7 Proof of Lemma 3.125. The situation with two triangles

Lemma 3.125 (Two-dimensional Case: Necessary Condition on the Triangulation) *Let $\Omega \subset \mathbb{R}^2$ be a polygonal domain and let $\{\mathcal{T}^h\}$ be a regular family of triangulations of Ω with triangles. Then the pair of spaces P_k/P_{k-1}, $k \geq 2$, does not satisfy the discrete inf-sup condition* (3.51) *if the triangulation consists of less than three triangles.*

Proof The cases of $\mathcal{T}^h$ consisting of one or two triangles will be studied separately.

$\mathcal{T}^h$ *consists of one triangle.* The discrete inf-sup condition is equivalent to the condition $\nabla \cdot V^h = Q^h$, see Lemma 3.12. This condition is equivalent to the satisfaction of the two conditions $\nabla \cdot V^h \subseteq Q^h$ and $Q^h \subseteq \nabla \cdot V^h$.

Consider first the reference mesh cell $\hat{K}$ and the vertex $\mathbf{0}$ at the origin. In this vertex, all derivatives of $\boldsymbol{v}^h \in V^h$ are determined by the homogeneous boundary conditions of V^h, i.e., all derivatives of $\boldsymbol{v}^h$ vanish and in particular it is $\nabla \cdot \boldsymbol{v}^h(\mathbf{0}) = 0$. However, the value $q^h(\mathbf{0})$ of an arbitrary $q^h \in Q^h$ is not necessarily zero and hence $Q^h \not\subseteq \nabla \cdot V^h$.

For a general mesh cell K consider the vertex $\boldsymbol{a}_1$ which is mapped to the origin. From (B.27) or (B.31), it follows that

$$\left(\nabla \cdot \boldsymbol{v}^h\right)(\boldsymbol{a}_1) = \left(DF_K^{-1}(F_K(\hat{\boldsymbol{x}}))\right)^T : D_{\hat{\boldsymbol{x}}}\hat{\boldsymbol{v}}^h(\mathbf{0}) = 0,$$

because all derivatives of $\hat{\boldsymbol{v}}^h$ vanish in the origin. With the same argument as before, it follows that $Q^h \not\subseteq \nabla \cdot V^h$.

$\mathcal{T}^h$ *consists of two triangles.* For this case, the situation as shown in Fig. 3.7 is considered. Similarly to the case with one triangle, the general situation can be mapped with suitable affine mappings of K_1 and K_2 to the situation depicted in Fig. 3.7.

The barycentric coordinates, see Definition B.31, are given by

$$\begin{aligned} &\lambda_{1,\boldsymbol{a}} = 1 + x - y,\ \lambda_{1,\boldsymbol{c}} = y, \\ &\lambda_{2,\boldsymbol{a}} = 1 - x - y,\ \lambda_{2,\boldsymbol{b}} = x,\ \lambda_{2,\boldsymbol{c}} = y. \end{aligned} \tag{3.129}$$

Let $L(x)$ be the Legendre polynomial of degree $(k-2)$ on $(0, 1)$ with respect to the weight $x(1-x)^3$, i.e., it holds

$$\int_0^1 x(1-x)^3 L(x) p_l(x)\, dx = 0 \tag{3.130}$$

for all polynomials of degree $l < k-2$. With the help of $L(x)$, an element from Q^h is defined which does not depend on y and for which it holds

$$\partial_x q^h(\boldsymbol{x}) = \begin{cases} -L(-x) & \text{if } x < 0, \\ L(x) & \text{if } x \geq 0. \end{cases} \tag{3.131}$$

It will be shown that this pressure is $L^2(\Omega)$ orthogonal $\nabla \cdot V^h$.

Considering the numerator in the discrete inf-sup condition (3.51), one finds with integration by parts that

$$\left(\nabla \cdot \boldsymbol{v}^h, q^h\right) = -\left(\boldsymbol{v}^h, \nabla q^h\right) = -\int_{K_1 \cup K_2} \left(v_1^h \partial_x q^h\right)(\boldsymbol{x})\, d\boldsymbol{x}. \tag{3.132}$$

Hence, only the first component of $\boldsymbol{v}^h$ needs to be taken into account. Since v_1^h belongs to the Taylor–Hood space, it has the following properties.

- It is on each mesh cell a polynomial of degree k.
- It vanishes at all edges of K_1 and K_2 which are not the common edge. This property can be imposed on each mesh cell by multiplying an arbitrary polynomial of degree $(k-2)$ with the appropriate barycentric coordinates, which are linear polynomials.
- It is continuous at $x = 0$. This property is achieved by decomposing the arbitrary polynomial of degree $(k-2)$ into an arbitrary polynomial of degree $(k-2)$ with respect to y and a polynomial of degree $(k-2)$ which vanishes at $x = 0$.

These requirements lead to the following general form of v_1^h

$$v_1^h(\boldsymbol{x}) = \begin{cases} \lambda_{1,\boldsymbol{a}}\lambda_{1,\boldsymbol{c}} \left(\hat{p}_{k-2}(y) + x\tilde{p}_{k-3}(x,y)\right) & \text{if } \boldsymbol{x} \in K_1, \\ \lambda_{2,\boldsymbol{a}}\lambda_{2,\boldsymbol{c}} \left(\hat{p}_{k-2}(y) + x\overline{p}_{k-3}(x,y)\right) & \text{if } \boldsymbol{x} \in K_2, \end{cases}$$

where the subscript indicates the degree of the polynomials. This expression can be inserted in (3.132), One gets for K_1 with (3.131), the substitution $\xi = -x$, the definition of the barycentric coordinates (3.129), and switching the lower and upper bound of the integral

$$\begin{aligned} &-\int_{K_1} \lambda_{1,\boldsymbol{a}}\lambda_{1,\boldsymbol{c}} \left(\hat{p}_{k-2}(y) + x\tilde{p}_{k-3}(x,y)\right) L(-x)\, d\boldsymbol{x} \\ &= -\int_{-1}^{0}\int_0^{1+x} (1+x-y)y \left(\hat{p}_{k-2}(y) + x\tilde{p}_{k-3}(x,y)\right) L(-x)\, dy\, dx \end{aligned}$$

$$= \int_1^0 \int_0^{1-\xi} (1-\xi-y)y\left(\hat{p}_{k-2}(y) - \xi\tilde{p}_{k-3}(-\xi,y)\right) L(\xi)\, dy\, d\xi$$
$$= -\int_{K_2} \lambda_{2,\boldsymbol{a}}\lambda_{2,\boldsymbol{c}} \left(\hat{p}_{k-2}(y) - x\tilde{p}_{k-3}(-x,y)\right) L(x)\, d\boldsymbol{x}.$$

Inserting this expression in (3.132), one obtains that the term with $\hat{p}_{k-2}(y)$ vanishes and one gets

$$\begin{aligned}\left(\nabla\cdot\boldsymbol{v}^h, q^h\right) &= -\int_{K_2} \lambda_{2,\boldsymbol{a}}\lambda_{2,\boldsymbol{b}}\lambda_{2,\boldsymbol{c}} \left(\overline{p}_{k-3}(x,y) - \tilde{p}_{k-3}(-x,y)\right) L(x)\, d\boldsymbol{x}\\ &= -\int_0^1 xL(x)\left(\int_0^{1-x}(1-x-y)y\breve{p}_{k-3}(x,y)dy\right)dx,\end{aligned} \qquad (3.133)$$

where $\breve{p}_{k-3}(x,y)$ is a polynomial of degree $(k-3)$. With a direct computation, one finds for $0 \le i+j \le k-3$

$$\begin{aligned}\int_0^{1-x}(1-x-y)yx^iy^j\, dy &= x^i(1-x)\frac{(1-x)^{j+2}}{j+2} - x^i\frac{(1-x)^{j+3}}{j+3}\\ &= (1-x)^3\frac{x^i(1-x)^j}{(j+2)(j+3)} = (1-x)^3 p_{k-3}(x),\end{aligned}$$

where $p_{k-3}(x)$ is a polynomial of degree $(k-3)$. Inserting this expression in (3.133) gives together with (3.130)

$$\left(\nabla\cdot\boldsymbol{v}^h, q^h\right) = -\int_0^1 xL(x)(1-x)^3 p_{k-3}(x)\, dx = 0.$$

Hence, the discrete inf-sup condition (3.51) cannot be satisfied. ■

Lemma 3.126 (Decomposition of the Domain) *Let $\Omega = \Omega_1 \cup \Omega_2$ and set*

$$V^h(\Omega_i) = \left\{\boldsymbol{v}^h \in V^h \ :\ \boldsymbol{v}^h = \boldsymbol{0} \text{ in } \Omega\setminus\Omega_i\right\}, \quad i = 1,2.$$

Suppose that the inf-sup conditions

$$\sup_{\boldsymbol{v}^h\in V^h(\Omega_1)\setminus\{\boldsymbol{0}\}} \frac{\int_{\Omega_1}\left(\nabla\cdot\boldsymbol{v}^h\right)q^h\, d\boldsymbol{x}}{\|\nabla\boldsymbol{v}^h\|_{L^2(\Omega_1)}} \ge \beta_1^h\left\|q^h\right\|_{L^2(\Omega_1)} \quad \forall\, q^h \in Q^h,$$
$$\sup_{\boldsymbol{v}^h\in V^h(\Omega_2)\setminus\{\boldsymbol{0}\}} \frac{\int_{\Omega_2}\left(\nabla\cdot\boldsymbol{v}^h\right)q^h\, d\boldsymbol{x}}{\|\nabla\boldsymbol{v}^h\|_{L^2(\Omega_2)}} \ge \beta_2^h\left\|q^h\right\|_{L^2(\Omega_2)} \quad \forall\, q^h \in Q^h,$$

hold. Then the inf-sup condition

$$\sup_{\boldsymbol{v}^h \in V^h(\Omega)\setminus\{\boldsymbol{0}\}} \frac{\left(\nabla \cdot \boldsymbol{v}^h, q^h\right)}{\|\nabla \boldsymbol{v}^h\|_{L^2(\Omega)}} \geq \beta_{\mathrm{is}}^h \left\|q^h\right\|_{L^2(\Omega)} \quad \forall\, q^h \in Q^h \tag{3.134}$$

is valid with $\beta_{\mathrm{is}}^h = \min\{\beta_1^h, \beta_2^h\}/2$. *If* Ω_1 *and* Ω_2 *are disjoint, then* $\beta_{\mathrm{is}}^h = \min\{\beta_1^h, \beta_2^h\}$.

Proof Let $q^h \in Q^h$ be given. From the satisfaction of the inf-sup condition in the subdomains, it follows from (3.46) that there is a function $\boldsymbol{v}_1^h \in V^h(\Omega_1)$ and a function $\boldsymbol{v}_2^h \in V^h(\Omega_1)$ such that

$$\int_{\Omega_i} \left(\nabla \cdot \boldsymbol{v}_i^h\right) q^h \, d\boldsymbol{x} = \left\|q^h\right\|_{L^2(\Omega_i)}^2, \quad \left\|\nabla \boldsymbol{v}_i^h\right\|_{L^2(\Omega_i)} \geq \frac{1}{\beta_i^h} \left\|q^h\right\|_{L^2(\Omega_i)}, \quad i = 1, 2. \tag{3.135}$$

The constant C in (3.46) is specified in the proof of Theorem 3.46.

Let $\boldsymbol{v}^h = \boldsymbol{v}_1^h + \boldsymbol{v}_2^h \in V^h$, then one gets with the first inequality of (3.135)

$$\begin{aligned}\left(\nabla \cdot \boldsymbol{v}^h, q^h\right) &= \int_{\Omega_1} \left(\nabla \cdot \boldsymbol{v}_1^h\right) q^h \, d\boldsymbol{x} + \int_{\Omega_2} \left(\nabla \cdot \boldsymbol{v}_2^h\right) q^h \, d\boldsymbol{x} \\ &= \left\|q^h\right\|_{L^2(\Omega_1)}^2 + \left\|q^h\right\|_{L^2(\Omega_2)}^2 \geq \left\|q^h\right\|_{L^2(\Omega)}^2. \end{aligned} \tag{3.136}$$

The triangle inequality, the second inequality of (3.135), and estimating the norms of the pressure in the subdomains by the norm in Ω from above yields

$$\begin{aligned}\left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)} &\leq \left\|\nabla \boldsymbol{v}_1^h\right\|_{L^2(\Omega_1)} + \left\|\nabla \boldsymbol{v}_2^h\right\|_{L^2(\Omega_2)} \leq \frac{1}{\beta_1^h} \left\|q^h\right\|_{L^2(\Omega_1)} + \frac{1}{\beta_2^h} \left\|q^h\right\|_{L^2(\Omega_2)} \\ &\leq \max\left\{\frac{1}{\beta_1^h}, \frac{1}{\beta_2^h}\right\} \left(\left\|q^h\right\|_{L^2(\Omega_1)} + \left\|q^h\right\|_{L^2(\Omega_2)}\right) \\ &\leq \frac{2}{\min\left\{\beta_1^h, \beta_2^h\right\}} \left\|q^h\right\|_{L^2(\Omega)}. \end{aligned} \tag{3.137}$$

Using (3.136) and (3.137), one obtains for all $q^h \in Q^h$ that there is a $\boldsymbol{v}^h \in V^h$ such that

$$\frac{\left(\nabla \cdot \boldsymbol{v}^h, q^h\right)}{\|\nabla \boldsymbol{v}^h\|_{L^2(\Omega)}} \geq \frac{1}{2} \min\left\{\beta_1^h, \beta_2^h\right\} \frac{\left\|q^h\right\|_{L^2(\Omega)}^2}{\|q^h\|_{L^2(\Omega)}} = \frac{1}{2} \min\left\{\beta_1^h, \beta_2^h\right\} \left\|q^h\right\|_{L^2(\Omega)}. \tag{3.138}$$

Taking the supremum over V^h on both sides of this inequality proves (3.134).

If the subdomains are disjoint, one does not need to estimate the norms of the pressure in the subdomains by the norm in Ω but one has with the second inequality of (3.135)

$$\begin{aligned}
\left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)}^2 &= \left\|\nabla \boldsymbol{v}_1^h\right\|_{L^2(\Omega_1)}^2 + \left\|\nabla \boldsymbol{v}_2^h\right\|_{L^2(\Omega_2)}^2 \\
&\le \left(\frac{1}{\beta_1^h}\right)^2 \left\|q^h\right\|_{L^2(\Omega_1)}^2 + \left(\frac{1}{\beta_2^h}\right)^2 \left\|q^h\right\|_{L^2(\Omega_2)}^2 \\
&\le \max\left\{\left(\frac{1}{\beta_1^h}\right)^2, \left(\frac{1}{\beta_2^h}\right)^2\right\} \left(\left\|q^h\right\|_{L^2(\Omega_1)}^2 + \left\|q^h\right\|_{L^2(\Omega_2)}^2\right) \\
&= \min\left\{\left(\beta_1^h\right)^2, \left(\beta_2^h\right)^2\right\} \left\|q^h\right\|_{L^2(\Omega)}^2 .
\end{aligned}$$

Taking the square root and applying the same reasoning as in (3.138) gives the last statement of the lemma. ∎

Remark 3.127 (More than Two Subdomains) If Ω is decomposed into more than two subdomains, a result which is similar to that of Lemma 3.126 can be proved. Estimate (3.136) is performed in the same way. In estimate (3.137), one gets a larger factor, depending on the number of overlappings of the subdomains. Consequently, one obtains a smaller factor in the corresponding global inf-sup condition.

The proof of the discrete inf-sup condition for the Taylor–Hood pair of spaces P_k/P_{k-1}, $k \ge 2$, starts with a decomposition of the triangulation into macroelements consisting of three adjacent triangles. It is remarked in Boffi et al. (2013, p. 499) that it is possible to prove that each triangulation of a polygonal domain can be represented as the disjoint union of triplets of adjacent triangles and of polygons than can be obtained as unions of triplets with at most three intersections. Hence, the maximal number of overlappings that have an impact on the constant in estimate (3.137) is independent of the triangulation. Consequently, the scaling factor of the global inf-sup constant is independent of the triangulation. □

Theorem 3.128 (Two-dimensional Case: Sufficient Condition on the Triangulation) *Let $\Omega \subset \mathbb{R}^2$ be a polygonal domain and let $\{\mathcal{T}^h\}$ be a regular family of triangulations of Ω with triangles. The pair of spaces P_k/P_{k-1}, $k \ge 2$, satisfies the discrete inf-sup condition* (3.51) *if the triangulation contains at least three triangles.*

Proof Let $\mathcal{T}^h$ be a member of the family of quasi-uniform triangulations of Ω. For proving the discrete inf-sup condition (3.51), it follows from Lemma 3.126 and Remark 3.127 that it is sufficient to prove the discrete inf-sup condition for an arbitrary triplet of triangles.

Consider an arbitrary triangle $K_2 \in \mathcal{T}^h$ with the adjacent triangles K_1 and K_3. The triangles are mapped with the reference map $F_{K_2}^{-1}$ such that K_2 becomes the reference triangle, see Fig. 3.8. For simplicity of notation, the triangles will be still denoted with K_1, K_2, K_3 and the domain spanned is $\overline{\omega} = K_1 \cup K_2 \cup K_3$.

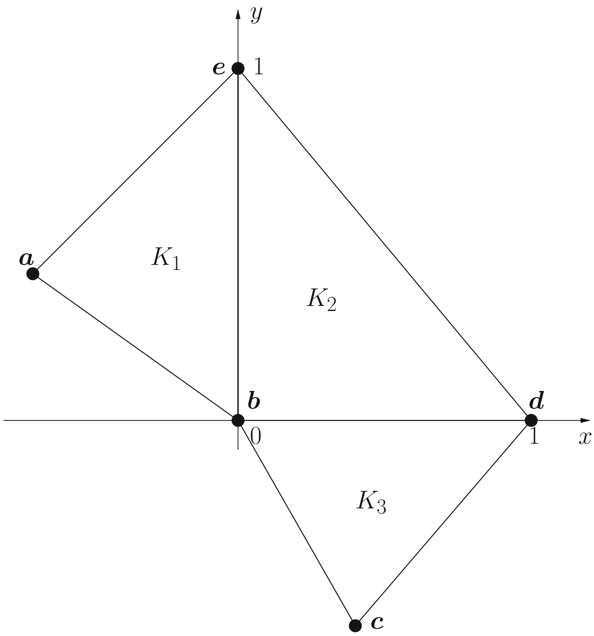

Fig. 3.8 Proof of Theorem 3.128. The situation with three triangles

The main idea for proving the discrete inf-sup condition for $V^h = V^h(\omega)$ and $Q^h = Q^h(\omega)$ consists in applying the criterion formulated in Theorem 3.94. To this end, it is sufficient to prove the conditions given in Corollary 3.95. The proof is performed by induction, starting with $k = 2$.

Preliminaries The barycentric coordinates in the mesh cell K_1 which vanish on the edge $E_{\boldsymbol{a},\boldsymbol{b}}$ are denoted by $\lambda_{1,\boldsymbol{a},\boldsymbol{b}}$ and so on.

Let $L_{1,l,x}(x)$ be the Legendre polynomial of degree l in $[x_{\boldsymbol{a}}, 0]$, with $x_{\boldsymbol{a}}$ being the x-coordinate of the vertex $\boldsymbol{a}$, with respect to the measure $\mu_{1,x}$, where the measure is defined by

$$\int_{x_{\boldsymbol{a}}}^{0} f(x)\, d\mu_{1,x} = \int_{K_1} \lambda_{1,\boldsymbol{a},\boldsymbol{b}}(x)\lambda_{1,\boldsymbol{a},\boldsymbol{e}}(x) f(x)\, d\boldsymbol{x}, \quad \forall f\ :\ [x_{\boldsymbol{a}}, 0] \to \mathbb{R}.$$

The Legendre polynomial satisfies

$$0 = \int_{x_{\boldsymbol{a}}}^{0} L_{1,l,x}(x) p_k(x)\, dx = \int_{K_1} \lambda_{1,\boldsymbol{a},\boldsymbol{b}}(x)\lambda_{1,\boldsymbol{a},\boldsymbol{e}}(x)_{1,l,x}(x) p_k(x)\, d\boldsymbol{x} \tag{3.139}$$

for all polynomials $p_k(x)$ with degree $k < l$. In the same way, other Legendre polynomials are defined: $L_{2,l,x}(x)$ with $\lambda_{2,\boldsymbol{d},\boldsymbol{e}}, \lambda_{2,\boldsymbol{b},\boldsymbol{d}}$, $L_{2,l,y}(y)$ with $\lambda_{2,\boldsymbol{d},\boldsymbol{e}}, \lambda_{2,\boldsymbol{b},\boldsymbol{e}}$, and

$L_{3,l,y}(y)$ with $\lambda_{3,\boldsymbol{b},\boldsymbol{c}}, \lambda_{3,\boldsymbol{c},\boldsymbol{d}}$. All Legendre polynomials are normalized such that they take the value 1 if the argument is 0.

Start of the Induction: The Case $k = 2$ Given an arbitrary $q^h \in Q^h$ then the function $\boldsymbol{v}^h = \left(v_1^h, v_2^h\right)^T$ is defined triangle by triangle in the following way

$$v_1^h(x,y) = \begin{cases} -\lambda_{1,\boldsymbol{a},\boldsymbol{b}}\lambda_{1,\boldsymbol{a},\boldsymbol{e}} \left\|\nabla q^h\right\|_{L^2(\omega)} \sigma & \text{in } K_1, \\ -\lambda_{2,\boldsymbol{d},\boldsymbol{e}}\lambda_{2,\boldsymbol{b},\boldsymbol{d}} \left\|\nabla q^h\right\|_{L^2(\omega)} \sigma - \lambda_{2,\boldsymbol{d},\boldsymbol{e}}\lambda_{2,\boldsymbol{b},\boldsymbol{e}}\partial_x q^h & \text{in } K_2, \\ -\lambda_{3,\boldsymbol{b},\boldsymbol{c}}\lambda_{3,\boldsymbol{c},\boldsymbol{d}}\partial_x q^h & \text{in } K_3, \end{cases}$$

$$v_2^h(x,y) = \begin{cases} -\lambda_{1,\boldsymbol{a},\boldsymbol{b}}\lambda_{1,\boldsymbol{a},\boldsymbol{e}}\partial_y q^h & \text{in } K_1, \\ -\lambda_{2,\boldsymbol{d},\boldsymbol{e}}\lambda_{2,\boldsymbol{b},\boldsymbol{e}} \left\|\nabla q^h\right\|_{L^2(\omega)} \tau - \lambda_{2,\boldsymbol{d},\boldsymbol{e}}\lambda_{2,\boldsymbol{b},\boldsymbol{d}}\partial_y q^h & \text{in } K_2, \\ -\lambda_{3,\boldsymbol{b},\boldsymbol{c}}\lambda_{3,\boldsymbol{c},\boldsymbol{d}} \left\|\nabla q^h\right\|_{L^2(\omega)} \tau & \text{in } K_3. \end{cases}$$

The factors σ and τ are chosen in $\{-1, 1\}$ such that the terms

$$\sigma_1 = \sigma \left\|\nabla q^h\right\|_{L^2(\omega)} \left(\int_{K_1} \lambda_{1,\boldsymbol{a},\boldsymbol{b}}\lambda_{1,\boldsymbol{a},\boldsymbol{e}}\partial_x q^h \, d\boldsymbol{x} + \int_{K_2} \lambda_{2,\boldsymbol{d},\boldsymbol{e}}\lambda_{2,\boldsymbol{b},\boldsymbol{d}}\partial_x q^h \, d\boldsymbol{x}\right),$$

$$\tau_1 = \tau \left\|\nabla q^h\right\|_{L^2(\omega)} \left(\int_{K_2} \lambda_{2,\boldsymbol{d},\boldsymbol{e}}\lambda_{2,\boldsymbol{b},\boldsymbol{e}}\partial_y q^h \, d\boldsymbol{x} + \int_{K_3} \lambda_{3,\boldsymbol{b},\boldsymbol{c}}\lambda_{3,\boldsymbol{c},\boldsymbol{d}}\partial_y q^h \, d\boldsymbol{x}\right)$$

are non-negative. The function $\boldsymbol{v}^h$ belongs to V^h:

- In each mesh cell, each term is the product of two barycentric coordinates, which are linear polynomials, and constant factors. Thus, $\boldsymbol{v}^h$ is in each mesh cell a polynomial of at most degree 2.
- The function $\boldsymbol{v}^h$ vanishes on $\partial\omega$ since in its definition the respective barycentric coordinates that vanish on $\partial\omega$ are involved.
- Finally, $\boldsymbol{v}^h$ is continuous at the edges $E_{\boldsymbol{b},\boldsymbol{e}}$ and $E_{\boldsymbol{b},\boldsymbol{d}}$. Consider for example $E_{\boldsymbol{b},\boldsymbol{e}}$, then it holds on this edge that $\lambda_{1,\boldsymbol{a},\boldsymbol{e}} = \lambda_{2,\boldsymbol{d},\boldsymbol{e}}$ and $\lambda_{1,\boldsymbol{a},\boldsymbol{b}} = \lambda_{2,\boldsymbol{b},\boldsymbol{d}}$. In addition, from the continuity of q^h it follows that the tangential derivative $\partial_x q^h$ is continuous. Hence, the continuity of $\boldsymbol{v}^h$ at $E_{\boldsymbol{b},\boldsymbol{e}}$ is proved. A similar reasoning can be applied for $E_{\boldsymbol{b},\boldsymbol{d}}$.

Now, the inequalities (3.99) have to be shown. Using that the barycentric coordinates take values in [0, 1] and the triangle inequality gives

$$\begin{aligned} &\left\|\boldsymbol{v}^h\right\|_{L^2(\omega)}^2 \\ &= \int_{K_1} \left(v_1^h\right)^2 + \left(v_2^h\right)^2 \, d\boldsymbol{x} + \int_{K_2} \left(v_1^h\right)^2 + \left(v_2^h\right)^2 \, d\boldsymbol{x} + \int_{K_3} \left(v_1^h\right)^2 + \left(v_2^h\right)^2 \, d\boldsymbol{x} \\ &\le \int_{K_1} \left(\left\|\nabla q^h\right\|_{L^2(\omega)}^2 + \left(\partial_y q^h\right)^2\right) \, d\boldsymbol{x} \end{aligned}$$

$$
\begin{aligned}
&+2\int_{K_2}\left(2\left\|\nabla q^h\right\|_{L^2(\omega)}^2+\left(\partial_x q^h\right)^2+\left(\partial_y q^h\right)^2\right)\,d\boldsymbol{x}\\
&+\int_{K_3}\left(\left(\partial_x q^h\right)^2+\left\|\nabla q^h\right\|_{L^2(\omega)}^2\right)\,d\boldsymbol{x}\\
\le{}& C\left[|\omega|\left\|\nabla q^h\right\|_{L^2(\omega)}^2+\left\|\nabla q^h\right\|_{L^2(K_1)}^2+\left\|\nabla q^h\right\|_{L^2(K_2)}^2+\left\|\nabla q^h\right\|_{L^2(K_3)}^2\right]\\
\le{}& C_2\left\|\nabla q^h\right\|_{L^2(\omega)}^2. \qquad (3.140)
\end{aligned}
$$

By this estimate it is obvious that the constant C_2 does not depend on q^h and $\boldsymbol{v}^h$.

To prove the other inequality in (3.99), one starts by showing that from $b\left(\boldsymbol{v}^h,q^h\right)=0$ it follows that $\nabla q^h=\mathbf{0}$. A straightforward calculation gives

$$
\begin{aligned}
0={}&-b\left(\boldsymbol{v}^h,q^h\right)=\int_\omega\left(\nabla\cdot\boldsymbol{v}^h\right)q^h\,d\boldsymbol{x}=-\int_\omega\boldsymbol{v}^h\nabla q^h\,d\boldsymbol{x}\\
={}&\int_{K_1}\lambda_{1,\boldsymbol{a},\boldsymbol{b}}\lambda_{1,\boldsymbol{a},\boldsymbol{e}}\left(\partial_y q^h\right)^2\,d\boldsymbol{x}+\sigma_1\\
&+\int_{K_2}\left(\lambda_{2,\boldsymbol{d},\boldsymbol{e}}\lambda_{2,\boldsymbol{b},\boldsymbol{e}}\left(\partial_x q^h\right)^2+\lambda_{2,\boldsymbol{d},\boldsymbol{e}}\lambda_{2,\boldsymbol{b},\boldsymbol{e}}\left(\partial_y q^h\right)^2\right)\,d\boldsymbol{x}\\
&+\tau_1+\int_{K_3}\lambda_{3,\boldsymbol{b},\boldsymbol{c}}\lambda_{3,\boldsymbol{c},\boldsymbol{d}}\left(\partial_x q^h\right)^2\,d\boldsymbol{x}. \qquad (3.141)
\end{aligned}
$$

Since the barycentric coordinates are positive in the interior of the mesh cells and σ_1 and τ_1 are defined to be non-negative, it follows that all terms have to vanish. One gets immediately $\partial_y q^h=0$ in K_1, $\partial_x q^h=\partial_y q^h=0$ in K_2, and $\partial_x q^h=0$ in K_3. From $\sigma_1=0$ and $\partial_x q^h=0$ in K_2, it follows that $\partial_x q^h=0$ in K_1. With the same argument, one obtains $\partial_y q^h=0$ in K_3. Altogether $\nabla q^h=\mathbf{0}$ in ω and there is a constant $C_1\left(q^h\right)$ such that

$$
b(\boldsymbol{v}^h,q^h)\ge C_1\left(q^h\right)\left\|\nabla q^h\right\|_{L^2(\omega)}^2.
$$

It has to be shown that one can choose a positive constant C_1 independent of q^h.

To this end, one observes first that considering the function αq^h with $\alpha\in\mathbb{R}\setminus\{0\}$, one obtains with the velocity $\alpha\boldsymbol{v}^h$ also an estimate with the constant $C_1\left(q^h\right)$. Hence, all possible values of $C_1\left(q^h\right)$ are taken on each sphere $S_q=\left\{q^h\in Q^h\ :\ \left\|q^h\right\|_{L^2(\Omega)}=q>0\right\}$. Since Q^h is a finite-dimensional space, S_q is a compact set.

On the other hand, considering the function $q_\varepsilon^h=q^h+\varepsilon\phi^h$ with arbitrary but fixed $\psi^h\in Q^h$, defining a function $\boldsymbol{v}_\varepsilon^h$ in the same spirit as $\boldsymbol{v}^h$, then one gets an estimate with a constant $C_1\left(q_\varepsilon^h\right)$. It is $\lim_{\varepsilon\to 0}q_\varepsilon^h=q^h$ and because of the continuity of the norm and the linearity of the differentiation one finds $\lim_{\varepsilon\to 0} \boldsymbol{v}_\varepsilon^h = \boldsymbol{v}^h$. It follows with the linearity of the integration and the continuity of the norm that

$$\lim_{\varepsilon\to 0} \frac{-b\left(\boldsymbol{v}_\varepsilon^h, q_\varepsilon^h\right)}{\left\|\nabla q_\varepsilon^h\right\|_{L^2(\omega)}^2} = \lim_{\varepsilon\to 0} \frac{-\int_\omega \boldsymbol{v}_\varepsilon^h \nabla q^h \, d\boldsymbol{x} - \varepsilon \int_\omega \boldsymbol{v}_\varepsilon^h \nabla \psi^h \, d\boldsymbol{x}}{\left\|\nabla q_\varepsilon^h\right\|_{L^2(\omega)}^2} = \frac{-\int_\omega \boldsymbol{v}^h \nabla q^h \, d\boldsymbol{x}}{\left\|\nabla q^h\right\|_{L^2(\omega)}^2} \geq C_1\left(q^h\right).$$

Hence, one can choose $C_1\left(q_\varepsilon^h\right)$ such that $C_1\left(q_\varepsilon^h\right) \to C_1\left(q^h\right)$ and the constants $C_1\left(q^h\right)$ can be chosen as a continuous function with respect to q^h.

Since $C_1\left(q^h\right)$ takes all of its values on S_q, $C_1\left(q^h\right)$ is continuous and S_q is compact, it follows from the Theorem of Weierstrass that $C_1\left(q^h\right)$ takes its infimum and this infimum is the minimum of $C_1\left(q^h\right)$. Because all values of $C_1\left(q^h\right)$ are positive, it follows now that

$$C_1 = \inf_{q^h \in S_q} C_1\left(q^h\right) = \min_{q^h \in S_q} C_1\left(q^h\right) > 0$$

is a positive constant that satisfies the first inequality in (3.99) independent of q^h.

Since the analysis was performed for the reference configuration depicted in Fig. 3.8, C_1 does not depend on the mesh width. Transforming this configuration to a configuration in a domain Ω requires the transformation of the numerator and of the denominator of the discrete inf-sup condition. Mapping the numerator gives the factor $|\det(B_K)|^{-1}$, where B_K is the matrix of the affine transform, see (B.35). Both terms of the denominator introduce the factor $|\det(B_K)|^{-1/2}$, compare (B.32) and (B.34). Hence, these factors cancel and the inf-sup constant for the configuration in Ω is independent of the mesh width.

Induction: The Case $k > 2$ Hypothesis of the induction: The statement of Theorem 3.128 is proved for P_{k-1}/P_{k-2}.

Consider again an arbitrary function $q^h \in Q^h$. It will be assumed that there is at least one triangle where q^h is exactly of degree $(k-1)$. Otherwise, the result follows from the hypothesis of the induction. A corresponding function $\boldsymbol{v}^h$ is defined in a similar way as for $k = 2$

$$v_1^h(x,y) = \begin{cases} -\lambda_{1,\boldsymbol{a},\boldsymbol{b}}\lambda_{1,\boldsymbol{a},\boldsymbol{e}} \left\|\nabla q^h\right\|_{L^2(\omega)} L_{1,k-2,x}\sigma & \text{in } K_1, \\ -\lambda_{2,\boldsymbol{d},\boldsymbol{e}}\lambda_{2,\boldsymbol{b},\boldsymbol{d}} \left\|\nabla q^h\right\|_{L^2(\omega)} L_{2,k-2,x}\sigma - \lambda_{2,\boldsymbol{d},\boldsymbol{e}}\lambda_{2,\boldsymbol{b},\boldsymbol{e}}\partial_x q^h & \text{in } K_2, \\ -\lambda_{3,\boldsymbol{b},\boldsymbol{c}}\lambda_{3,\boldsymbol{c},\boldsymbol{d}}\partial_x q^h & \text{in } K_3, \end{cases}$$

$$v_2^h(x,y) = \begin{cases} -\lambda_{1,\boldsymbol{a},\boldsymbol{b}}\lambda_{1,\boldsymbol{a},\boldsymbol{e}}\partial_y q^h & \text{in } K_1, \\ -\lambda_{2,\boldsymbol{d},\boldsymbol{e}}\lambda_{2,\boldsymbol{b},\boldsymbol{e}} \left\|\nabla q^h\right\|_{L^2(\omega)} L_{2,k-2,y}\tau - \lambda_{2,\boldsymbol{d},\boldsymbol{e}}\lambda_{2,\boldsymbol{b},\boldsymbol{d}}\partial_y q^h & \text{in } K_2, \\ -\lambda_{3,\boldsymbol{b},\boldsymbol{c}}\lambda_{3,\boldsymbol{c},\boldsymbol{d}} \left\|\nabla q^h\right\|_{L^2(\omega)} L_{3,k-2,y}\tau & \text{in } K_3, \end{cases}$$

with the factors σ and τ chosen in $\{-1, 1\}$ such that the terms

$$\sigma_1 = \sigma \left\|\nabla q^h\right\|_{L^2(\omega)} \left(\int_{K_1} \lambda_{1,\boldsymbol{a},\boldsymbol{b}} \lambda_{1,\boldsymbol{a},\boldsymbol{e}} L_{1,k-2,x} \partial_x q^h \, d\boldsymbol{x} + \int_{K_2} \lambda_{2,\boldsymbol{d},\boldsymbol{e}} \lambda_{2,\boldsymbol{b},\boldsymbol{d}} L_{2,k-2,x} \partial_x q^h \, d\boldsymbol{x} \right),$$

$$\tau_1 = \tau \left\|\nabla q^h\right\|_{L^2(\omega)} \left(\int_{K_2} \lambda_{2,\boldsymbol{d},\boldsymbol{e}} \lambda_{2,\boldsymbol{b},\boldsymbol{e}} L_{2,k-2,y} \partial_y q^h \, d\boldsymbol{x} + \int_{K_3} \lambda_{3,\boldsymbol{b},\boldsymbol{c}} \lambda_{3,\boldsymbol{c},\boldsymbol{d}} L_{3,k-2,y} \partial_y q^h \, d\boldsymbol{x} \right)$$

are non-negative. Using similar arguments as in the case $k = 2$, one shows that $\boldsymbol{v}^h \in V^h$. The additional arguments are the followings:

- The restriction of $\boldsymbol{v}^h$ to a mesh cell is at most a polynomial of degree k since the Legendre polynomials are polynomials of degree $(k - 2)$ and the barycentric coordinates are polynomials of degree 1.
- Because of the normalization, the Legendre polynomials $L_{1,k-2,x}$ and $L_{2,k-2,x}$ take the same values at $x = 0$. Likewise $L_{2,k-2,y}$ and $L_{3,k-2,y}$ take the same value at $y = 0$. These properties ensure the continuity of $\boldsymbol{v}^h$.

The proof of the second estimate of (3.99) is performed in the same way as (3.140), using in addition that the Legendre polynomials are bounded from above, since they are continuous functions, and these bounds do not depend on q^h and $\boldsymbol{v}^h$.

Assuming now that $b\left(\boldsymbol{v}^h, q^h\right) = 0$, one finds with an analog of the calculation (3.141) and with the same arguments as for $k = 2$ that $\partial_y q^h = 0$ in K_1, $\partial_x q^h = \partial_y q^h = 0$ in K_2, $\partial_x q^h = 0$ in K_3, $\sigma_1 = \tau_1 = 0$. From $\sigma_1 = 0$ and $\partial_x q^h = 0$ in K_2, one gets

$$\int_{K_1} \lambda_{1,\boldsymbol{a},\boldsymbol{b}} \lambda_{1,\boldsymbol{a},\boldsymbol{e}} L_{1,k-2,x} \partial_x q^h \, d\boldsymbol{x} = 0.$$

The property (3.139) gives that $\partial_x q^h$ is in K_1 a polynomial of degree less than $(k-2)$. It follows that q^h is in K_1 a polynomial of degree less than $(k - 1)$. In the same way, one finds that q^h is in K_3 a polynomial of degree less than $(k - 1)$. Altogether, q^h is in each mesh cell a polynomial of degree less than $(k - 1)$, in contradiction to the assumption. Hence there is a constant $C_1\left(q^h\right)$ such that

$$b\left(\boldsymbol{v}^h, q^h\right) \geq C_1\left(q^h\right) \left\|\nabla q^h\right\|_{L^2(\omega)}^2.$$

The arguments for finding a constant C_1 independent of q^h and h are the same as in the case $k = 2$. ■

Theorem 3.129 (Three-dimensional Case) *Let $\Omega \subset \mathbb{R}^3$ be a polygonal domain and let $\{\mathcal{T}^h\}$ be a regular family of decompositions of Ω in tetrahedra. It is assumed that every tetrahedron has at least one interior vertex. Then, the pair of spaces P_k/P_{k-1}, $k \geq 2$, satisfies the discrete inf-sup condition* (3.51)*.*

Proof The proof is performed with the macroelement technique introduced in Sect. 3.5.4.

Consider an overlapping macroelement partition of $\mathcal{T}^h$ and define for each vertex $\boldsymbol{a}$ in the interior of Ω the macroelement

$$M_{\boldsymbol{a}} = \left\{K \; : \; K \in \mathcal{T}^h, \boldsymbol{a} \text{ is vertex of } K\right\}.$$

Since a regular family of triangulations is considered, there is a uniform upper bound for the number of tetrahedra sharing a common vertex. Hence, there is uniform finite number of types of macroelements. From the assumption that each tetrahedron has a vertex in the interior of Ω, it follows that each tetrahedron belongs to at least on macroelement. Each mesh cell belongs to at most four macroelements. Hence, the geometric assumptions of Lemma 3.109 and Remark 3.112 are satisfied. It remains to show the macroelement condition (3.110).

Let $K \in M_{\boldsymbol{a}}$ and let E be an edge of K where $\boldsymbol{a}$ is one of its vertices. The coordinate system is chosen such that E is on the x-axis. Take any function $q^h \in Q^h_{\text{grad},M} \cap Q^h_{0,M}$. It will be shown first that q^h is constant on K. To this end, define the function

$$\boldsymbol{v}^h = -\left(\lambda_{i,1}\lambda_{i,2}\partial_x q^h, 0, 0\right)^T \quad \text{in } K_i,$$

where $\{K_i\}$ is the set of mesh cells in $M_{\boldsymbol{a}}$ sharing the edge E. The functions $\lambda_{i,1}$ and $\lambda_{i,2}$ are the barycentric coordinates of K_i with respect to the two faces of K_i that do not share the edge E. On all other mesh cells of $M_{\boldsymbol{a}}$, it is set $\boldsymbol{v}^h = \boldsymbol{0}$. The function $\boldsymbol{v}^h$ belongs to the space $V^h_{0,M}$ for the following reasons.

- The function $\boldsymbol{v}^h$ is a polynomial of degree k since $\partial_x q^h$ is a polynomial of degree $(k-2)$ and $\lambda_{i,1}$, $\lambda_{i,2}$ are both linear polynomials.
- By construction with the appropriate barycentric coordinates it follows that $\boldsymbol{v}^h|_{\partial M_{\boldsymbol{a}}} = \boldsymbol{0}$.
- Since q^h is continuous, the tangential derivatives of q^h on all faces are continuous. The edge E is a tangential vector on all faces with this edge, thus $\partial_x q^h$ is continuous on all faces with edge E. Likewise, the barycentric coordinates are continuous on these faces. On all other faces, $\boldsymbol{v}^h$ vanishes. Altogether, $\boldsymbol{v}^h$ is continuous.

Since $q^h \in Q^h_{\text{grad},M}$, one finds

$$0 = \left(\nabla \cdot \boldsymbol{v}^h, q^h\right)_M = -\left(\boldsymbol{v}^h, \nabla q^h\right)_M = \sum_{K_i \in M_{\boldsymbol{a}}} \int_{K_i} \lambda_{i,1}\lambda_{i,2}\left(\partial_x q^h\right)^2 \, d\boldsymbol{x}.$$

Consequently $\partial_x q^h = 0$ and q^h does not depend on x in K_i, in particular not in K. Applying the same argument for the two other edges of K with vertex $\boldsymbol{a}$, one finds that three partial derivatives of q^h vanish in K. Since the edges form linearly independent directions, it follows that $\nabla q^h = \boldsymbol{0}$ in K.

The same procedure can be applied for all $K \in M_{\boldsymbol{a}}$. It follows that q^h is a constant in each mesh cell in $M_{\boldsymbol{a}}$ and since q^h is continuous, one gets that q^h is a constant. Because the only constant in $Q^h_{0,M}$ is zero, it follows that $q^h = 0$, which is just the macroelement condition (3.110). ■

Remark 3.130 (Q_k/Q_{k-1}, $k \geq 2$) The discrete inf-sup condition (3.51) for the pair Q_2/Q_1 in two dimensions was proved in Stenberg (1984) using the macroelement technique developed in this paper. The macroelement consists of two mesh cells. The analysis was extended to higher order pairs of Q_k/Q_{k-1} spaces in Stenberg (1990). For three dimensions, the discrete inf-sup condition was proved for Q_2/Q_1 in Stenberg (1987). □

Remark 3.131 (The Modified Taylor–Hood Finite Element) Let the triangulation $\mathcal{T}^h$ consist of triangles, or of quadrilaterals, or of hexahedra. These mesh cells are regularly refined, e.g., by connecting the barycenters of the edges of a triangle, leading to a triangulation $\mathcal{T}^{h/2}$.

Considering for simplicity triangles, then the modified Taylor–Hood finite element spaces are defined by using as velocity space the $P_1^{h/2}$ finite element space on the fine grid and as pressure space the P_1 finite element space on the coarse grid. The number of degrees of freedom of $P_1^{h/2}$ is the same as of the space P_2 on the coarse grid, but instead of piecewise quadratic functions, piecewise linear functions are considered. Similar constructions can be performed for quadrilaterals and hexahedra. A reason for using the modified Taylor–Hood space might be that the available code only supports lowest order finite elements.

The satisfaction of the discrete inf-sup condition for the modified Taylor–Hood pair of spaces can be proved, e.g., with the technique from Bercovier and Pironneau (1979) and Verfürth (1984) which is mentioned in Remark 3.124. □

3.6.3 Spaces on Simplicial Meshes with Discontinuous Pressure

Remark 3.132 (Motivation) A motivation for using finite element spaces with discontinuous pressure found in the literature is that such spaces satisfy a local mass conservation. This aspect is discussed in Remark 4.32. □

Remark 3.133 (Basic Ideas for the Construction) The pair of finite element spaces $P_2^{\text{bubble}}/P_1^{\text{disc}}$ proposed in Crouzeix and Raviart (1973), the Scott–Vogelius pair of finite element spaces, P_k/P_{k-1}, $k \geq 2$, introduced in Scott and Vogelius (1985), and the Bernardi–Raugel spaces from Bernardi and Raugel (1985) apply discontinuous pressure approximations on simplicial grids. It was already shown in Examples 3.70

and 3.73 that using the combination $P_k/P_{k-1}^{\text{disc}}$ generally violates the discrete inf-sup condition (3.51).

The Scott–Vogelius finite element considers still $P_k/P_{k-1}^{\text{disc}}$, $k \geq d$, but on special meshes, which allow to show the satisfaction of the discrete inf-sup condition. The idea of the $P_2^{\text{bubble}}/P_1^{\text{disc}}$ finite element and the Bernardi–Raugel finite element consists in enriching the discrete velocity space with bubble functions until it is sufficiently large such that the discrete inf-sup condition is satisfied, compare Remark 3.52. □

Remark 3.134 (The Scott–Vogelius Pair of Finite Element Spaces) This pair of finite element spaces is given by $P_k/P_{k-1}^{\text{disc}}$, $k \geq 2$. Since

$$\nabla \cdot V^h = \nabla \cdot P_k = P_{k-1}^{\text{disc}} = Q^h,$$

finite element velocities from this pair are weakly divergence-free, which is a desirable property. However, as already demonstrated in Example 3.73, the Scott–Vogelius finite element generally does not satisfy the discrete inf-sup condition (3.51). But it can be proved that the pair $P_k/P_{k-1}^{\text{disc}}$ satisfies the discrete inf-sup condition in special situations, i.e., on special meshes. In this section, an overview on the known results will be given. □

Remark 3.135 ($P_k/P_{k-1}^{\text{disc}}$ in Two Dimensions) The fulfillment of the discrete inf-sup condition (3.51) was proved already in Scott and Vogelius (1985) in the two-dimensional case for $k \geq 4$ if there is no so-called singular vertex in the mesh. An internal vertex is said to be singular if edges which meet at the vertex fall onto two straight lines. In the counterexample for the general fulfillment of the discrete inf-sup condition, a singular vertex was considered, see Fig. 3.5.

The basic idea to overcome this problem consists in using meshes without singular vertices. To this end, so-called barycentric-refined grids are constructed. Starting from any admissible triangular mesh, new edges are introduced by connecting all vertices of a mesh cell with the barycenter of this mesh cell. This step creates smaller triangles, see Fig. 3.9 for an example. On barycentric-refined meshes, the $P_k/P_{k-1}^{\text{disc}}$, $k \in \{2, 3\}$, pair of finite element spaces was shown to satisfy the discrete inf-sup condition in Qin (1994), see also Example 4.144 for a proof in the case $k = 2$. Note that the case $k \geq 4$ is covered by the analysis from Scott and Vogelius (1985).

The use of the P_2/P_1^{disc} pair of finite element spaces on barycentric-refined meshes can be found occasionally in the literature, in particular to demonstrate the advantages of using pairs of finite element spaces which provide weakly divergence-free velocity solutions, e.g., see John et al. (2015) and the references therein. □

Remark 3.136 ($P_k/P_{k-1}^{\text{disc}}$ in Three Dimensions) In three dimensions, the use of barycentric-refined meshes avoids singular vertices and singular edges. In Zhang (2005), it was shown that the pair $P_k/P_{k-1}^{\text{disc}}$, $k \geq 3$, satisfies the discrete inf-sup condition on such meshes. The proof uses the macroelement technique from Stenberg (1984) introduced in Sect. 3.5.4.

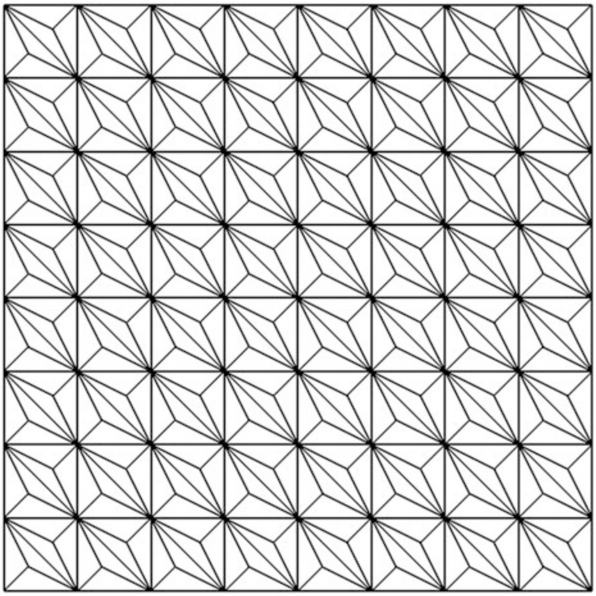

Fig. 3.9 Barycentric-refined simplicial grid on the unit square

The Scott–Vogelius pair of finite element spaces P_3/P_2^{disc} on barycentric-refined meshes in three dimensions was used in electro-chemical applications, see Fuhrmann et al. (2009, 2011), and in the simulation of turbulent flows in Cuff et al. (2015). □

Remark 3.137 (The $P_2^{\text{bubble}}/P_1^{\text{disc}}$ and $P_3^{\text{bubble}}/P_2^{\text{disc}}$ Pairs of Spaces) This pair was proposed in Crouzeix and Raviart (1973). To get an inf-sup stable pair of finite element spaces with a piecewise linear but discontinuous pressure on arbitrary simplicial grids, one has to enrich the velocity space. It turned out that it suffices to use bubble functions, i.e., local functions. In the two-dimensional case, one needs to add just all mesh cell bubble functions. For each triangular mesh cell, the cell bubble function is given by the product of the barycentric coordinates $\lambda_1\lambda_2\lambda_3$. In three dimensions, the enrichment proposed in Crouzeix and Raviart (1973) is given by all mesh cell bubble functions and all face bubble functions. For a mesh cell K with barycentric coordinates $\lambda_1, \ldots, \lambda_4$, a mesh cell bubble function is given by $\lambda_1\lambda_2\lambda_3\lambda_4$ and the four face bubble functions are the product of three mutually distinct barycentric coordinates, e.g., $\lambda_1\lambda_2\lambda_3$. Thus, for each velocity component, there are 15 degrees of freedom such that the dimension of the local space for enrichment is 45. From the practical point of view one has to take into consideration that bubbles are higher degree polynomials and their use requires the application of higher order quadrature rules. The $P_2^{\text{bubble}}/P_1^{\text{disc}}$ pair of spaces is considered to be a good one in Gresho and Sani (2000, p. 553).

The approach of stabilizing pairs of finite element spaces by enrichment with bubble functions can be extended to higher order spaces. For instance, in Crouzeix and Raviart (1973), one finds the pair $P_3^{\text{bubble}}/P_2^{\text{disc}}$ in two dimensions. □

Remark 3.138 (The Bernardi–Raugel Element of First Order P_1^{BR}/P_0) In the construction of the Bernardi–Raugel finite element of first order, the space $P_1(K)$ is enriched with genuinely vector-valued basis functions. Let λ_i, $i = 1, \ldots, d+1$, be the barycentric coordinates given in Definition B.31. With these functions, face (or in two dimensions edge) bubble functions are defined by

$$\boldsymbol{v}_{\text{bub},i}^h = \prod_{j=1, j\neq i}^{d+1} \lambda_j \boldsymbol{n}_i, \quad i = 1, \ldots, d+1,$$

where $\boldsymbol{n}_i$ is the unit outward pointing normal on the face E_i which is opposite to the vertex $\boldsymbol{a}_i$. The face bubble functions are polynomials. The local finite element space is given by

$$P_1^{\text{BR}}(K) = P_1(K) \oplus \text{span}\left\{\boldsymbol{v}_{\text{bub},i}^h, i = 1, \ldots, d+1\right\}.$$

The degrees of freedom are the values at the vertices $\boldsymbol{v}^h(\boldsymbol{a}_i)$, $i = 1, \ldots, d+1$, ($d(d+1)$ values) and the fluxes through the faces $\boldsymbol{v}^h \cdot \boldsymbol{n}_i$, $i = 1, \ldots, d+1$, ($(d+1)$ values). Hence, the number of local degrees of freedom used for enrichment is $(d+1)^2$.

The corresponding global velocity space consists of continuous functions. The face bubble functions of two mesh cells K_1 and K_2 with the common face E_i are continuous. By using for both mesh cells the outward pointing normal on E, one gets $\boldsymbol{v}_{\text{bub},i}^h(K_1)\big|_{E_i} = -\,\boldsymbol{v}_{\text{bub},i}^h(K_2)\big|_{E_i}$. But also the flux across the face has an opposite sign for this definition of the normals, such that the coefficients in front of $\boldsymbol{v}_{\text{bub},i}^h(K_1)$ and $\boldsymbol{v}_{\text{bub},i}^h(K_2)$ have the same absolute value but the opposite sign. Hence, the proposed combination of the basis functions and the local functionals leads to a globally continuous function.

The proof of the discrete inf-sup condition for the global space is based on the construction of a Fortin operator, see Lemma 3.78. This operator has the form (3.87) with $P_1^h = P_{\text{Cle}}^h$, see also Remark 3.83. It is defined, for each mesh cell K, by

$$\begin{aligned} P_{\text{For}}^h \boldsymbol{v}\,(\boldsymbol{a}_i) &= P_{\text{Cle}}^h\,(\boldsymbol{a}_i)\,, \\ \int_{E_i} \left(\boldsymbol{v} - P_{\text{For}}^h \boldsymbol{v}\right) \cdot \boldsymbol{n}_i\, ds &= 0, \quad i = 1, \ldots, d+1. \end{aligned}$$

The properties (3.82) from the Fortin criterion are proved in Bernardi and Raugel (1985, Lemma II.4).

The pair P_1^{BR}/P_0 is used only occasionally for academic purposes. □

Remark 3.139 (Families of Triangulations with Quadrilateral Mesh Cells) It is remarked in Bernardi and Raugel (1985) that both, the construction of the velocity space and the proof of the inf-sup condition, can be extended to regular families of triangulations consisting of quadrilaterals. This result was used in proving the discrete inf-sup condition for spaces with discontinuous pressure on quadrilateral grids, see Remarks 3.143 and 3.145. □

Remark 3.140 (The Bernardi–Raugel Element of Order Two $P_2^{\mathrm{BR}}/P_1^{\mathrm{disc}}$ in Three Dimensions) On tetrahedral grids, the space

$$P_2^{\mathrm{BR}}(K) = P_2(K) \oplus \mathrm{span}\left\{\boldsymbol{v}_{\mathrm{bub},i}^h, i = 1, \ldots, 4\right\} \oplus \left(\mathrm{span}\left\{\lambda_1\lambda_2\lambda_3\lambda_4\right\}\right)^3$$

was proposed for the local velocity space in Bernardi and Raugel (1985). This proposal, in connection with the discontinuous pressure space, can be considered as the extension of a pair of spaces introduced in Crouzeix and Raviart (1973) for two dimensions. The degrees of freedom of $P_2^{\mathrm{BR}}(K)$ are the values at the vertices and at the midpoints of the edges (from P_2: 30 degrees of freedom), the flux across the faces (from span $\{\boldsymbol{v}_{\mathrm{bub},i}\}$: four degrees of freedom), and the moments $\int_K x_i \nabla \cdot \boldsymbol{v}^h \, d\boldsymbol{x}$, $i = 1, 2, 3$. Thus, the dimension of the local space is 37.

With the same arguments as in Remark 3.138, one can show that the corresponding global velocity space consists of continuous functions. The satisfaction of the discrete inf-sup condition (3.51) can be proved with the construction of a Fortin operator, in a similar way as sketched in Remark 3.138.

Numerical simulations with the $P_2^{\mathrm{BR}}/P_1^{\mathrm{disc}}$ pair of finite element spaces can be found in John et al. (2010). In this paper, the main reason given for the preference of this pair to the Taylor–Hood pair P_2/P_1 is that the used multigrid solver, see Sect. 9.2.2, shows a better efficiency for pairs of spaces with discontinuous pressure. Altogether, this pair of spaces seems to be used so far only for academic purposes. □

3.6.4 Spaces on Quadrilateral and Hexahedral Meshes with Discontinuous Pressure

Remark 3.141 (The Spaces $Q_k/P_{k-1}^{\mathrm{disc}}$) The most common pairs of spaces with conforming velocity and discontinuous pressure on quadrilateral and hexahedral meshes are the spaces $Q_k/P_{k-1}^{\mathrm{disc}}$, $k \geq 2$. It was already shown in Example 3.71 that $Q_1/P_0 = Q_1/Q_0$ is in general not inf-sup stable. For $k \geq 2$, one has to distinguish two cases, the so-called mapped and the unmapped $Q_k/P_{k-1}^{\mathrm{disc}}$ spaces.

In the unmapped case, the local space $Q_k(K)$ is defined by a mapping from a reference cell $\hat{K}$ but the space $P_{k-1}^{\mathrm{disc}}(K)$ is defined directly on the mesh cell K. The mapped version defines both spaces with the reference transformation. Since the reference transformation from a quadrilateral or hexahedral reference cell is in general a bilinear or trilinear mapping, it gives rise to mesh cells with curved boundaries. In addition, in general it does not preserve the type of mapped functions, i.e., the images of polynomials are in general not polynomials. Thus, the mapped and unmapped version of $Q_k/P_{k-1}^{\text{disc}}$ are generally different on arbitrary meshes.

All simulations with $Q_k/P_{k-1}^{\text{disc}}$, $k \geq 2$, presented in this monograph were performed with the mapped version. □

Remark 3.142 (On the Interpolation Error of Mapped $Q_k/P_{k-1}^{\text{disc}}$ Pairs of Spaces) Since the mapped functions are no longer polynomials for arbitrary mesh cells, the application of the Bramble–Hilbert lemma, see Lemma C.5, is affected and optimal interpolation estimates cannot be proved, as shown in Arnold et al. (2002). In the same paper, examples of numerical simulations are presented where the corresponding reduction of the order of convergence can be observed.

In Arnold et al. (2002), Matthies (2001) optimal interpolation error estimates for mapped finite elements on quadrilaterals and hexahedra were studied. It turned out that the optimality is given for special families of triangulations. In two and three dimensions, families of meshes, which are obtained by a regular uniform refinement of an initial coarse grid, are among these special families. □

Remark 3.143 (The Unmapped $Q_k/P_{k-1}^{\text{disc}}$ Pairs of Spaces in Two Dimensions) The proof of the fulfillment of the discrete inf-sup condition (3.51) for the unmapped $Q_k/P_{k-1}^{\text{disc}}$, $k \geq 2$, finite element spaces in two dimensions can be found in Girault and Raviart (1986, Chap. II, Theorem. 3.2). It uses the macroelement technique from Boland and Nicolaides (1983), see Theorem 3.101. The application of this theorem requires the choice of an appropriate space $\overline{V}^h \subset V^h$. In Girault and Raviart (1986), the Bernardi–Raugel space of first order for quadrilateral mesh cells was chosen, see Remark 3.139. □

Example 3.144 (The Discrete Inf-Sup Condition for the Unmapped Q_2/P_1^{disc} Finite Element in Two Dimensions) The satisfaction of the discrete inf-sup condition (3.51) for the unmapped Q_2/P_1^{disc} finite element in two dimensions can be proved also easily with the macroelement technique of Stenberg (1984), as it was shown in this paper.

One has to prove that the macroelement condition, see Definition 3.107, is satisfied. To this end, take as macroelement just one mesh cell. Thus, there is just one class of macroelements and the macroelement decomposition of the domain coincides with its triangulation. The pressure is linear on the macroelement

$$Q_M^h = \left\{q^h \ : \ q^h\big|_M = a + b_x x + b_y y, \quad a, b_x, b_y \in \mathbb{R}\right\}.$$

The only degrees of freedom of $V_{0,M}^h$ are at the interior node. Taking the basis function $\boldsymbol{\phi}(\boldsymbol{x}) = (\varphi(\boldsymbol{x}), 0)^T$ which has the value $(1, 0)$ in this node, one obtains

$$\left(\nabla \cdot \boldsymbol{v}^h, q^h\right)_M = -\left(\boldsymbol{v}^h, \nabla q^h\right)_M = -b_x \int_M \varphi(\boldsymbol{x})\, d\boldsymbol{x}.$$

Since the integral of the basis function is positive, the condition $0 = \left(\nabla \cdot \boldsymbol{v}^h, q^h\right)_M$ implies that $b_x = 0$. Applying the basis function $(0, \varphi(\boldsymbol{x}))$ leads in the same way to

$b_y = 0$. It follows that

$$Q^h_{\mathrm{grad},M} = \left\{q^h \ : \ q^h\big|_M = a, \quad a \in \mathbb{R}\right\}$$

and consequently condition (3.110) is also satisfied. □

Remark 3.145 (The Mapped $Q_k/P_{k-1}^{\mathrm{disc}}$ Pairs of Spaces) The discrete inf-sup condition (3.51) for the mapped $Q_k/P_{k-1}^{\mathrm{disc}}$ pairs of spaces was proved for any space dimension d in Matthies and Tobiska (2002). The proof relies on the macroelement technique of Boland and Nicolaides (1983), see Theorem 3.101. For the construction of the subspace, which is required in the assumptions of this lemma, the proof of the discrete inf-sup condition for the extension of the two-dimensional Bernardi–Raugel element from triangles to quadrilaterals, mentioned in Remark 3.139, was extended to quadrilateral-type mesh cells of any space dimension. □

3.6.5 Non-conforming Finite Element Spaces of Lowest Order

Remark 3.146 (Non-conforming Finite Element Spaces) In the most general sense, non-conforming finite element methods are all methods where the finite element space is not a subspace of the function space used in the variational problem. This property might be caused, e.g., if for a problem, the domain Ω with curvilinear parts of the boundary is approximated by a domain Ω^h with polygonal or polyhedral boundary. But usually, one speaks of non-conforming finite element methods only if the non-inclusion of the spaces comes from the construction of the finite element space and it is independent of the special problem.

For incompressible flow models, the consideration of non-conforming discretizations will allow to define pairs of lowest order finite element spaces that satisfy the discrete inf-sup condition (3.51). The non-conformity is present only for the velocity but not for the pressure, i.e., $V^h \not\subset V$ and $Q^h \subset Q$.

Here, it will be concentrated on lowest order non-conforming discretizations because these are the most important non-conforming methods for incompressible flow problems. On simplicial meshes, this discretization is the so-called Crouzeix–Raviart finite element P_1^{nc}/P_0. That means, the velocity is approximated by a piecewise linear function that is continuous at the barycenters of the faces of the mesh cells, see Example B.43 for a detailed description, and the pressure is approximated by a piecewise constant function, see Example B.37.

The extension of this approach to quadrilateral and hexahedral meshes is the Rannacher–Turek element Q_1^{rot}/Q_0. For this element, the velocity approximation is achieved by rotated d-linear functions that have continuous degrees of freedom on the faces of the mesh cells, see Example B.53. The pressure is discretized by a piecewise constant function, see Example B.49.

Besides the possibility of using lowest order spaces, non-conforming finite element of lowest order possess some additional advantages. They can be used for the construction of efficient multigrid solvers or preconditioners for higher order discretizations of incompressible flow problems, see Sect. 9.2.2. Implementing the code for solving the Navier–Stokes equations on parallel computers, non-conforming discretizations generally require less communication overhead than conforming finite element methods. However, non-conforming finite elements are often more complicated from the point of view of numerical analysis.

The discrete inf-sup condition is proved with the construction of an operator $V \to V^h$ such that for each function $\boldsymbol{v} \in V$ analogs to the conditions (3.82) of a Fortin operator are satisfied. □

Remark 3.147 (The Non-conforming Q_1^{rot}/Q_0 Pairs of Finite Element Spaces) These pairs of finite element spaces were introduced and analyzed in Rannacher and Turek (1992) and later studied comprehensively in Schieweck (1997). They are defined on quadrilateral or hexahedral mesh cells. The local velocity spaces, called rotated bilinear finite element spaces, are defined as follows

$$\begin{aligned} Q_1^{\text{rot}}\left(\hat{K}\right) &= \left\{\hat{p}\ :\ \hat{p} \in \operatorname{span}\{1, \hat{x}_1, \hat{x}_2, \hat{x}_1^2 - \hat{x}_2^2\}\right\} && \text{if } d = 2, \\ Q_1^{\text{rot}}\left(\hat{K}\right) &= \left\{\hat{p}\ :\ \hat{p} \in \operatorname{span}\{1, \hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_1^2 - \hat{x}_2^2, \hat{x}_2^2 - \hat{x}_3^2\}\right\} && \text{if } d = 3, \\ Q_1^{\text{rot}}(K) &= \left\{p = \hat{p} \circ F_K^{-1}\ :\ \hat{p} \in Q_1^{\text{rot}}\left(\hat{K}\right)\right\}, \end{aligned}$$

where $\hat{K} = [-1, 1]^d$. There are two types of the Q_1^{rot} spaces, depending on the choice of the local functionals. Let $\mathcal{E}(K)$ be the set of all $(d-1)$-dimensional faces of K and let K' be a neigbor cell of K. Then, the mean-value-oriented finite element space Q_1^{rot} is given by

$$\begin{aligned} Q_1^{\text{rot}} = \Big\{ & v \in L^2(\Omega)\ :\ v|_K \in Q_1^{\text{rot}}(K), \\ & \int_E v|_K(\boldsymbol{s})\, d\boldsymbol{s} = \int_E v|_{K'}(\boldsymbol{s})\, d\boldsymbol{s} \ \ \forall\, E \in \mathcal{E}(K) \cap \mathcal{E}\left(K'\right)\Big\}, \end{aligned} \tag{3.142}$$

and the point-value-oriented finite element space Q_1^{rot} by

$$\begin{aligned} Q_1^{\text{rot}} = \{ & v \in L^2(\Omega)\ :\ v|_K \in Q_1^{\text{rot}}(K), v \text{ is continuous at the barycenters} \\ & \text{of the faces of } K\}. \end{aligned}$$

The mean-value-oriented and the point-value-oriented Q_1^{rot} are in general not equivalent.

The pressure is approximated by a piecewise constant function. □

Remark 3.148 (Extension of the Bilinear Form $b(\cdot,\cdot)$) For non-conforming spaces $V^h \not\subset V$, the bilinear form $b(\cdot,\cdot)$ from (3.34) is not defined for all $\boldsymbol{v}^h \in V^h$. Its domain of definition has to be extended appropriately. Since also the non-conforming finite element functions are piecewise polynomials, a natural approach consists in considering the bilinear form on each mesh cell and to define the extension by the sum over all mesh cells. Hence, for non-conforming finite element approximations, the bilinear form $b^h(\cdot,\cdot)$ as given in (3.53) has to be used. □

Remark 3.149 (Imposing Dirichlet Boundary Conditions) For the point-value-oriented non-conforming finite element spaces, the value of the Dirichlet boundary condition in the barycenter of the faces at the boundary is taken. Using the mean-value-oriented spaces, one computes the integrals of the boundary condition on these faces and normalizes with the area of the faces to set the boundary values. In the case of homogeneous Dirichlet boundary conditions, the boundary values computed in both ways are zero. □

Lemma 3.150 (A Norm in P_1^{nc} and Q_1^{rot}) *Consider the case of homogeneous Dirichlet boundary conditions. Then, the map $(a^h(\cdot,\cdot))^{1/2}$ with*

$$a^h\left(v^h, w^h\right) = \sum_{K\in\mathcal{T}^h} \left(\nabla v^h, \nabla w^h\right)_K$$

defines a norm in P_1^{nc} and Q_1^{rot}.

Proof One has to check the properties of Definition A.6.

i). If $v^h = 0$ in Ω, then

$$\left(a^h\left(v^h, v^h\right)\right)^{1/2} = \left(\sum_{K\in\mathcal{T}^h} \left(\nabla v^h, \nabla v^h\right)_K\right)^{1/2} = \left(\sum_{K\in\mathcal{T}^h} (\mathbf{0},\mathbf{0})_K\right)^{1/2} = 0.$$

Let $\left(a^h\left(v^h, v^h\right)\right)^{1/2} = 0$. Then v^h is a piecewise constant function, since the gradient in each mesh cell vanishes. The value of this constant in a mesh cell is known if the value of v^h in one point of the mesh cell is known.

The degrees of freedom of the P_1^{nc} and the point value-oriented Q_1^{rot} finite element functions are the values in the barycenter of the faces. On the boundary, these are just the values of the boundary condition, i.e., these values are zero. Hence, the constant is zero on all mesh cells, which have a face at the boundary. Consequently, the value is zero in all other barycenters of faces of these mesh cells. Since the finite element functions are continuous in the barycenters of the faces, the constant is zero also in the neighboring mesh cells. By induction, one obtains that the constant is zero in all mesh cells, i.e., $v^h = 0$ in Ω.

For the mean value-oriented Q_1^{rot} finite element, the integral on the faces which belong to Γ must vanish because of the homogeneous Dirichlet boundary condition. The only constant which fulfills this requirement is zero. That means, the constant is zero on all mesh cells with faces on the boundary. It follows that the integral on the other faces of these mesh cell is zero, too. By the definition of the mean value-oriented Q_1^{rot} finite element (3.142), the integrals on these faces of the functions from the neighboring mesh cells must also vanish. Since v^h is also constant on these mesh cells, the constant must be always zero. Again, it follows by induction that $v^h = 0$ in Ω.

ii). This property follows by a straightforward calculation

$$\left(a^h\left(\alpha v^h,\alpha v^h\right)\right)^{1/2} = \left(\sum_{K\in\mathcal{T}^h}\left(\nabla\left(\alpha v^h\right),\nabla\left(\alpha v^h\right)\right)_K\right)^{1/2}$$
$$= |\alpha|\left(\sum_{K\in\mathcal{T}^h}\left(\nabla v^h,\nabla v^h\right)_K\right)^{1/2} = |\alpha|\left(a^h\left(v^h,v^h\right)\right)^{1/2}$$

for all $\alpha \in \mathbb{R}$.

iii). The proof of the triangle inequality uses the triangle inequality on the individual mesh cells and the Cauchy–Schwarz inequality for sums (A.2)

$$\begin{aligned}
&\left(a^h\left(v^h+w^h,v^h+w^h\right)\right)^{1/2} \\
&= \left(\sum_{K\in\mathcal{T}^h}\left\|\nabla(v^h+w^h)\right\|_{L_2(K)}^2\right)^{1/2} \\
&\le \left(\sum_{K\in\mathcal{T}^h}\left(\left\|\nabla v^h\right\|_{L_2(K)}+\left\|\nabla w^h\right\|_{L_2(K)}\right)^2\right)^{1/2} \\
&= \left(\sum_{K\in\mathcal{T}^h}\left\|\nabla v^h\right\|_{L_2(K)}^2 + 2\left\|\nabla v^h\right\|_{L_2(K)}\left\|\nabla w^h\right\|_{L_2(K)} + \left\|\nabla w^h\right\|_{L_2(K)}^2\right)^{1/2} \\
&\le \left(\sum_{K\in\mathcal{T}^h}\left\|\nabla v^h\right\|_{L_2(K)}^2 + 2\left(\sum_{K\in\mathcal{T}^h}\left\|\nabla v^h\right\|_{L_2(K)}^2\right)^{1/2}\left(\sum_{K\in\mathcal{T}^h}\left\|\nabla w^h\right\|_{L_2(K)}^2\right)^{1/2}\right. \\
&\qquad\left. + \sum_{K\in\mathcal{T}^h}\left\|\nabla w^h\right\|_{L_2(K)}^2\right)^{1/2} \\
&= \left\{\left[\left(\sum_{K\in\mathcal{T}^h}\left\|\nabla v^h\right\|_{L_2(K)}^2\right)^{1/2} + \left(\sum_{K\in\mathcal{T}^h}\left\|\nabla w^h\right\|_{L_2(K)}^2\right)^{1/2}\right]^2\right\}^{1/2} \\
&= \left(\sum_{K\in\mathcal{T}^h}\left\|\nabla v^h\right\|_{L_2(K)}^2\right)^{1/2} + \left(\sum_{K\in\mathcal{T}^h}\left\|\nabla w^h\right\|_{L_2(K)}^2\right)^{1/2} \\
&= \left(a^h\left(v^h,v^h\right)\right)^{1/2} + \left(a^h\left(w^h,w^h\right)\right)^{1/2}
\end{aligned}$$

for all $v^h, w^h \in V^h$. ■

Theorem 3.151 (Inf-Sup Condition for the Lowest Order Crouzeix–Raviart Finite Element P_1^{nc}/P_0) *The lowest order Crouzeix–Raviart pair of finite element spaces P_1^{nc}/P_0 satisfies the discrete inf-sup condition* (3.51) *or equivalently* (3.52)*. The constant β_{is}^h in* (3.51) *can be chosen as $\beta_{\text{is}}^h = \beta_{\text{is}}$, where β_{is} is the constant of the continuous inf-sup condition* (3.14)*.*

Proof The proof uses the properties that the functions from P_1^{nc} are piecewise linear and the functions from P_0 are piecewise constant.

Let $\boldsymbol{v} \in V$ be an arbitrary function. In the first step of the proof, an appropriate interpolation of $\boldsymbol{v}$ to a function $P_E^h\boldsymbol{v} \in V^h = P_1^{\text{nc}}$ is constructed. The functions from P_1^{nc} are determined by their values in the barycenters of the faces of the mesh cells. Let $\left\{\overline{\mathcal{E}}^h\right\}$ be the set of faces, $\{\boldsymbol{m}_E\}$ be the corresponding set of barycenters, and $\left\{\mathcal{E}^h\right\}$ be the set of all interior faces, then the values of $P_E^h\boldsymbol{v}$ at $\{\boldsymbol{m}_E\}$ are defined as follows

$$P_E^h\boldsymbol{v}\,(\boldsymbol{m}_E) = \begin{cases} \dfrac{1}{|E|}\displaystyle\int_E \boldsymbol{v}(\boldsymbol{s})\,d\boldsymbol{s} & \text{if } E \in \mathcal{E}^h, \\ \boldsymbol{0} & \text{if } E \in \overline{\mathcal{E}}^h \setminus \mathcal{E}^h. \end{cases} \tag{3.143}$$

This construction is just an interpolation using the global functionals of the definition of the finite element space, see (B.20) for the space P_1^{nc}. These functionals are continuous.

In the second step of the proof, a relation between $b^h\left(P_E^h\boldsymbol{v}, q^h\right)$ and $b\left(\boldsymbol{v}, q^h\right)$ is shown for arbitrary $q^h \in Q^h$. Using that q^h is constant and integration by parts, one obtains

$$\begin{aligned} b^h\left(P_E^h\boldsymbol{v}, q^h\right) &= -\sum_{K\in\mathcal{T}^h}\int_K \left(\nabla\cdot P_E^h\boldsymbol{v}\; q^h\right)(\boldsymbol{x})\,d\boldsymbol{x} \\ &= -\sum_{K\in\mathcal{T}^h} q^h\big|_K \int_K \nabla\cdot P_E^h\boldsymbol{v}(\boldsymbol{x})\,d\boldsymbol{x} \\ &= -\sum_{K\in\mathcal{T}^h} q^h\big|_K \left(\sum_{E\subset\partial K}\int_E \left(P_E^h\boldsymbol{v}\cdot\boldsymbol{n}_E\right)(\boldsymbol{s})\,d\boldsymbol{s}\right). \end{aligned}$$

Since $\left(P_E^h\boldsymbol{v}\cdot\boldsymbol{n}_E\right)\big|_E$ is linear, the integral is computed exactly with the mid point rule, such that

$$\begin{aligned} b^h\left(P_E^h\boldsymbol{v}, q^h\right) &= -\sum_{K\in\mathcal{T}^h} q^h\big|_K \left(\sum_{E\subset\partial K} |E|\,P_E^h\boldsymbol{v}(\boldsymbol{m}_E)\cdot\boldsymbol{n}_E\right) \\ &= -\sum_{K\in\mathcal{T}^h} q^h\big|_K \left(\sum_{E\subset\partial K}\int_E (\boldsymbol{v}\cdot\boldsymbol{n}_E)(\boldsymbol{s})\,d\boldsymbol{s}\right) \end{aligned}$$

$$
\begin{aligned}
&= - \sum_{K \in \mathcal{T}^h} q^h\big|_K \int_K \nabla \cdot \boldsymbol{v}(\boldsymbol{x})\, d\boldsymbol{x} \\
&= - \sum_{K \in \mathcal{T}^h} \int_K \left(\nabla \cdot \boldsymbol{v}\, q^h\right)(\boldsymbol{x})\, d\boldsymbol{x} \\
&= b^h\left(\boldsymbol{v}, q^h\right) = b\left(\boldsymbol{v}, q^h\right),
\end{aligned}
$$

where in the second step the definition of $P_E^h \boldsymbol{v}(\boldsymbol{m}_E)$ was used and then again integration by parts was applied. Altogether, this step gives

$$
b^h\left(P_E^h \boldsymbol{v}, q^h\right) = b\left(\boldsymbol{v}, q^h\right) \quad \forall\ \boldsymbol{v} \in V,\ \forall\ q^h \in Q^h. \tag{3.144}
$$

This equation is the analog of the first condition for a Fortin operator in (3.82).

The next step of the proof bounds $\left\| P_E^h \boldsymbol{v} \right\|_{V^h}$ by $\|\boldsymbol{v}\|_V$. Let $K \in \mathcal{T}^h$ be an arbitrary mesh cell, then one obtains with integration by parts, using that $\Delta P_E^h \boldsymbol{v}\big|_K = \boldsymbol{0}$ and $\left(\nabla P_E^h \boldsymbol{v} \cdot \boldsymbol{n}_E\right)\big|_E$ is constant,

$$
\begin{aligned}
\left\| \nabla P_E^h \boldsymbol{v} \right\|_{L^2(K)}^2 &= \int_K \left(\nabla P_E^h \boldsymbol{v} : \nabla P_E^h \boldsymbol{v}\right)(\boldsymbol{x})\, d\boldsymbol{x} \\
&= \sum_{E \subset \partial K} \int_E \left(P_E^h \boldsymbol{v} \cdot \nabla P_E^h \boldsymbol{v} \boldsymbol{n}_E\right)(\boldsymbol{s})\, d\boldsymbol{s} - \int_K \left(P_E^h \boldsymbol{v} \cdot \Delta P_E^h \boldsymbol{v}\right)(\boldsymbol{x})\, d\boldsymbol{x} \\
&= \sum_{E \subset \partial K} \left(\nabla P_E^h \boldsymbol{v} \boldsymbol{n}_E\right)\big|_E \cdot \int_E P_E^h \boldsymbol{v}(\boldsymbol{s})\, d\boldsymbol{s}.
\end{aligned}
$$

Since the function in the integral is linear, the midpoint rule of integration computes its exact value. Applying then the definition (3.143) of $P_E^h \boldsymbol{v}\ (\boldsymbol{m}_E)$, again integration by parts, $\Delta P_E^h \boldsymbol{v}\big|_K = \boldsymbol{0}$, and finally the Cauchy–Schwarz inequality (A.10), one gets

$$
\begin{aligned}
\left\| \nabla P_E^h \boldsymbol{v} \right\|_{L^2(K)}^2 &= \sum_{E \subset \partial K} \left(\nabla P_E^h \boldsymbol{v} \boldsymbol{n}_E\right)\big|_E \cdot |E|\, P_E^h \boldsymbol{v}(\boldsymbol{m}_E) \\
&= \sum_{E \subset \partial K} \left(\nabla P_E^h \boldsymbol{v} \boldsymbol{n}_E\right)\big|_E \cdot \int_E \boldsymbol{v}(\boldsymbol{s})\, d\boldsymbol{s} \\
&= \sum_{E \subset \partial K} \int_E \left(\boldsymbol{v} \cdot \nabla P_E^h \boldsymbol{v} \boldsymbol{n}_E\right)(\boldsymbol{s})\, d\boldsymbol{s} \\
&= \int_K \left(\nabla P_E^h \boldsymbol{v} : \nabla \boldsymbol{v}\right)(\boldsymbol{x})\, d\boldsymbol{x} + \int_K \left(\boldsymbol{v} \cdot \Delta P_E^h \boldsymbol{v}\right)(\boldsymbol{x})\, d\boldsymbol{x} \\
&\le \left\| \nabla P_E^h \boldsymbol{v} \right\|_{L^2(K)} \|\nabla \boldsymbol{v}\|_{L^2(K)}.
\end{aligned}
$$

Using this estimate and the Cauchy–Schwarz inequality for sums (A.2) yields

$$\begin{aligned}\left\|P_E^h\boldsymbol{v}\right\|_{V^h}^2 &= \sum_{K\in\mathcal{T}^h}\left\|\nabla P_E^h\boldsymbol{v}\right\|_{L^2(K)}^2\\ &\le \sum_{K\in\mathcal{T}^h}\left\|\nabla P_E^h\boldsymbol{v}\right\|_{L^2(K)}\left\|\nabla\boldsymbol{v}\right\|_{L^2(K)} \le \left\|P_E^h\boldsymbol{v}\right\|_{V^h}\left\|\boldsymbol{v}\right\|_V.\end{aligned}$$

The result of this step is that

$$\left\|P_E^h\boldsymbol{v}\right\|_{V^h} \le \left\|\boldsymbol{v}\right\|_V \quad \forall\ \boldsymbol{v}\in V. \tag{3.145}$$

This estimate corresponds to the second condition for a Fortin operator in (3.82).

From (3.144) and (3.145), one gets

$$\frac{b\left(\boldsymbol{v},q^h\right)}{\left\|\boldsymbol{v}\right\|_V} \le \frac{b^h\left(P_E^h\boldsymbol{v},q^h\right)}{\left\|P_E^h\boldsymbol{v}\right\|_{V^h}} \quad \forall\ \boldsymbol{v}\in V, \forall\ q^h\in Q^h.$$

Next, applying the inf-sup condition (3.14) of the continuous problem yields, since $Q^h\subset Q$,

$$\begin{aligned}\beta_{\text{is}}\left\|q^h\right\|_Q &\le \sup_{\boldsymbol{v}\in V,\boldsymbol{v}\ne\boldsymbol{0}}\frac{b\left(\boldsymbol{v},q^h\right)}{\left\|\boldsymbol{v}\right\|_V} \le \sup_{\boldsymbol{v}\in V,\boldsymbol{v}\ne\boldsymbol{0}}\frac{b^h\left(P_E^h\boldsymbol{v},q^h\right)}{\left\|P_E^h\boldsymbol{v}\right\|_{V^h}}\\ &\le \sup_{\boldsymbol{v}^h\in V^h,\boldsymbol{v}^h\ne\boldsymbol{0}}\frac{b^h\left(\boldsymbol{v}^h,q^h\right)}{\left\|\boldsymbol{v}^h\right\|_{V^h}} \quad \forall\ q^h\in Q^h,\end{aligned}$$

where the final estimate uses that the set of projections is a subset of V^h. The supremum in this subset cannot be larger than the supremum in V^h. This inequality is exactly the discrete inf-sup condition (3.52) with $\beta_{\text{is}}^h = \beta_{\text{is}}$. ■

Remark 3.152 (Anisotropic Mesh Cells) Note that the proof of Theorem 3.151 does not impose any special requirement on the aspect ratio of the simplicial mesh cells. In addition, the discrete inf-sup constant equals the continuous one, thus the discrete inf-sup constant does not depend on the mesh. Consequently, even for grids with anisotropic mesh cells, i.e., mesh cells K where h_K/ρ_K is very large, the discrete inf-sup condition is satisfied uniformly with $\beta_{\text{is}}^h = \beta_{\text{is}}$ and one has a well-posed finite element problem. □

Remark 3.153 (The Discrete Inf-Sup Condition for the Q_1^{rot}/Q_0 Finite Element) The fulfillment of the discrete inf-sup condition (3.51) was proved in Rannacher and Turek (1992) with the construction of a Fortin operator, see Lemma 3.78. To this end, the global interpolation operators to Q_1^{rot} were used, which are generated by the canonical local interpolation operators to $Q_1^{\text{rot}}(K)$. For the mean-value-oriented version, the proof of the properties (3.82) of the Fortin operator is similar to the proof of Theorem 3.151.

A comprehensive analysis for several realizations of the Q_1^{rot}/Q_0 pair of finite element spaces can be found in Schieweck (1997). This analysis relaxes the assumptions on the meshes for proving the inf-sup condition compared with the analysis from Rannacher and Turek (1992). □

Remark 3.154 (The Unmapped Version of the Q_1^{rot}/Q_0 Finite Element Spaces) There are also unmapped (non-parametric) versions of these finite element spaces, which define the polynomials directly on the mesh cell K. It is shown in Rannacher and Turek (1992) that these versions are inf-sup stable on more general meshes than the mapped (parametric) version of the Q_1^{rot}/Q_0 finite element, e.g., on strongly nonuniform meshes. Considering all four types of Q_1^{rot}/Q_0 finite elements, the optimal order of convergence on perturbed meshes is achieved only by the mean-value-oriented version of the unmapped Q_1^{rot}/Q_0 finite element. □

Remark 3.155 (Practical Use of the Q_1^{rot}/Q_0 Finite Element Spaces) In the simulations presented in this monograph, the mapped mean-value-oriented Q_1^{rot} finite element space was used in two dimensions and the mapped point-value-oriented Q_1^{rot} finite element space in three dimensions. For $d = 3$, the integrals on the faces of mesh cells, whose equality is required in the mapped mean-value-oriented Q_1^{rot} finite element space, involve a weighting function which depends on the particular mesh cell K. The computation of these weighting functions for all mesh cells is an additional computational overhead such that Schieweck (1997, p. 21) recommended to use for $d = 3$ the simpler mapped point-value-oriented form of the Q_1^{rot} finite element. □

Remark 3.156 (Non-conforming Finite Element Spaces of Higher Order) The paper Matthies and Tobiska (2005) presents a generalization of the P_1^{nc}/P_0 pair of spaces in two dimensions to arbitrary order. The satisfaction of the discrete inf-sup condition for the proposed pairs of spaces is proved. □

3.6.6 Computing the Discrete Inf-Sup Constant

Remark 3.157 (The Constant in the Discrete Inf-Sup Condition) The actual size of the constant β_{is}^h in the discrete inf-sup condition (3.51) is of some interest since it enters, e.g., finite element error estimates. Of course, this constant depends on the concrete spaces, i.e., in particular on the domain and on the used mesh. □

Remark 3.158 (On the Supremum) Consider a fixed $q^h \in Q^h$, then $b^h\left(\boldsymbol{v}^h, q^h\right)$, defined in (3.53), is a linear functional in $\left(V^h\right)'$. The linearity follows from the linearity of integration and the boundedness follows with the application of the triangle inequality, the Cauchy–Schwarz inequality (A.10), the Cauchy–Schwarz inequality for sums (A.2), and (3.40) for $\Omega = K$

$$\begin{aligned}\left|b^h\left(\boldsymbol{v}^h, q^h\right)\right| &\le \sum_{K\in\mathcal{T}^h} \left|\left(\nabla\cdot\boldsymbol{v}^h, q^h\right)_K\right| \le \sum_{K\in\mathcal{T}^h} \left\|\nabla\cdot\boldsymbol{v}^h\right\|_{L^2(K)} \left\|q^h\right\|_{L^2(K)} \\ &\le \left(\sum_{K\in\mathcal{T}^h} \left\|\nabla\cdot\boldsymbol{v}^h\right\|_{L^2(K)}^2\right)^{1/2} \left(\sum_{K\in\mathcal{T}^h} \left\|q^h\right\|_{L^2(K)}^2\right)^{1/2} \\ &\le C\left(\sum_{K\in\mathcal{T}^h} \left\|\nabla\boldsymbol{v}^h\right\|_{L^2(K)}^2\right)^{1/2} \left\|q^h\right\|_{L^2(\Omega)} = C\left\|\boldsymbol{v}^h\right\|_{V^h}.\end{aligned}$$

Due to the Representation Theorem of Riesz, Theorem B.3, there is an element $\boldsymbol{v}_b^h \in V^h$ such that

$$\left(\boldsymbol{v}_b^h, \boldsymbol{v}^h\right)_{V^h} = b^h\left(\boldsymbol{v}^h, q^h\right) \quad \forall\ \boldsymbol{v}^h \in V^h. \tag{3.146}$$

It follows that

$$b^h\left(\boldsymbol{v}^h, q^h\right) = \left(\boldsymbol{v}_b^h, \boldsymbol{v}^h\right)_{V^h} \le \left\|\boldsymbol{v}_b^h\right\|_{V^h} \left\|\boldsymbol{v}^h\right\|_{V^h} \quad \forall\ \boldsymbol{v}^h \in V^h,$$

such that one obtains an upper bound for the supremum

$$\sup_{\boldsymbol{v}^h\in V^h\setminus\{\boldsymbol{0}\}} \frac{b^h\left(\boldsymbol{v}^h, q^h\right)}{\left\|\boldsymbol{v}^h\right\|_{V^h}} \le \sup_{\boldsymbol{v}^h\in V^h\setminus\{\boldsymbol{0}\}} \frac{\left\|\boldsymbol{v}_b^h\right\|_{V^h}\left\|\boldsymbol{v}^h\right\|_{V^h}}{\left\|\boldsymbol{v}^h\right\|_{V^h}} = \left\|\boldsymbol{v}_b^h\right\|_{V^h}. \tag{3.147}$$

The upper bound is even attained. Choosing in particular $\boldsymbol{v}^h = \boldsymbol{v}_b^h$ yields, with (3.146),

$$\frac{b^h\left(\boldsymbol{v}_b^h, q^h\right)}{\left\|\boldsymbol{v}_b^h\right\|_{V^h}} = \frac{\left\|\boldsymbol{v}_b^h\right\|_{V^h}^2}{\left\|\boldsymbol{v}_b^h\right\|_{V^h}} = \left\|\boldsymbol{v}_b^h\right\|_{V^h}. \tag{3.148}$$

Combining (3.147) and (3.148) shows that

$$\boldsymbol{v}_b^h = \operatorname*{arg\,sup}_{\boldsymbol{v}^h\in V^h\setminus\{\boldsymbol{0}\}} \frac{b^h\left(\boldsymbol{v}^h, q^h\right)}{\left\|\boldsymbol{v}^h\right\|_{V^h}}.$$

With these considerations, one finds that the discrete inf-sup constant is given by

$$\left(\beta_{\mathrm{is}}^h\right)^2 = \inf_{q^h\in Q^h\setminus\{0\}} \sup_{\boldsymbol{v}^h\in V^h\setminus\{\boldsymbol{0}\}} \frac{\left(b^h\left(\boldsymbol{v}^h, q^h\right)\right)^2}{\left\|\boldsymbol{v}^h\right\|_{V^h}^2 \left\|q^h\right\|_Q^2} = \inf_{q^h\in Q^h\setminus\{0\}} \frac{\left\|\boldsymbol{v}_b^h\right\|_{V^h}^2}{\left\|q^h\right\|_Q^2}. \tag{3.149}$$

Up to this point, one can argue analogously for the continuous inf-sup condition. □

Remark 3.159 (Generalized Eigenvalue Problem for Computing the Value of the Discrete Inf-Sup Constant) Let $\{\boldsymbol{\phi}_i^h\}_{i=1}^{N_u}$ and $\{\psi_i^h\}_{i=1}^{N_p}$ be bases of V^h and Q^h, respectively. Then, one can define the Gramian matrices or mass matrices

$$M_V = \left(\left(\boldsymbol{\phi}_j^h, \boldsymbol{\phi}_i^h\right)_{V^h}\right)_{i,j=1}^{N_u}, \quad M_Q = \left(\left(\psi_j^h, \psi_i^h\right)_{Q^h}\right)_{i,j=1}^{N_p}$$

and the matrix

$$B = b^h\left(\boldsymbol{\phi}_j^h, \psi_i^h\right), \quad i = 1, \ldots, N_p, j = 1, \ldots, N_u.$$

With the bases, one has also representations of the finite element functions in terms of vectors

$$\boldsymbol{v}^h = \sum_{i=1}^{N_u} v_i \boldsymbol{\phi}_i^h, \quad \boldsymbol{v}_b^h = \sum_{i=1}^{N_u} b_i \boldsymbol{\phi}_i^h, \quad q^h = \sum_{i=1}^{N_p} q_i \psi_i^h,$$

with the vectors $\underline{v}$, $\underline{b}$, and $\underline{q}$, respectively. The inner products in V^h and Q^h and the bilinear form $b^h(\cdot,\cdot)$ can be written now in terms of matrices and vectors, e.g.,

$$\left(\boldsymbol{v}_b^h, \boldsymbol{v}^h\right)_{V^h} = \underline{b}^T M_V^T \underline{v}, \quad b^h\left(\boldsymbol{v}^h, q^h\right) = \underline{v}^T B^T \underline{q}, \quad \left\|\boldsymbol{v}_b^h\right\|_{V^h}^2 = \underline{b}^T M_V^T \underline{b}.$$

In this way, one obtains for (3.146)

$$\underline{b}^T M_V^T \underline{v} = \underline{v}^T B^T \underline{q} \quad \Longleftrightarrow \quad \underline{v}^T M_V \underline{b} = \underline{v}^T B^T \underline{q} \quad \forall\, \underline{v} \in \mathbb{R}^{N_u},$$

from what follows that

$$M_V \underline{b} = B^T \underline{q} \quad \Longrightarrow \quad \underline{b} = M_V^{-1} B^T \underline{q}.$$

Inserting this expression in (3.149) yields

$$\begin{aligned}\left(\beta_{\text{is}}^h\right)^2 &= \inf_{\underline{q}\in\mathbb{R}^{N_q}\setminus\{\underline{0}\}} \frac{\underline{b}^T M_V^T \underline{b}}{\underline{q}^T M_Q^T \underline{q}} = \inf_{\underline{q}\in\mathbb{R}^{N_q}\setminus\{\underline{0}\}} \frac{\underline{q}^T B M_V^{-T} M_V^T M_V^{-1} B^T \underline{q}}{\underline{q}^T M_Q^T \underline{q}} \\ &= \inf_{\underline{q}\in\mathbb{R}^{N_q}\setminus\{\underline{0}\}} \frac{\underline{q}^T B M_V^{-1} B^T \underline{q}}{\underline{q}^T M_Q^T \underline{q}}.\end{aligned}$$

Using that M_V and M_Q are symmetric and positive definite, one can rewrite the right-hand side in the form

$$\frac{\left(\underline{q}^T M_Q^{1/2}\right)\left(M_Q^{-1/2} B M_V^{-T/2}\right)\left(M_V^{-1/2} B^T M_Q^{-T/2}\right)\left(M_Q^{T/2}\underline{q}\right)}{\left(\underline{q}^T M_Q^{1/2}\right)\left(M_Q^{T/2}\underline{q}\right)},$$

which shows that this expression is a Rayleigh quotient. It is known, see Lemma A.19, that the infimum of the Rayleigh quotient is attained and it is the smallest eigenvalue of the eigenvalue problem

$$\left(M_Q^{-1/2} B M_V^{-T/2}\right)\left(M_V^{-1/2} B^T M_Q^{-T/2}\right)\left(M_Q^{T/2} \underline{q}\right) = \lambda\left(M_Q^{T/2} \underline{q}\right).$$

Multiplying this problem with $M_Q^{1/2}$ shows that it is equivalent to computing the smallest eigenvalue of the generalized eigenvalue problem

$$B M_V^{-1} B^T \underline{q} = \lambda M_Q^T \underline{q}. \tag{3.150}$$

The square root of this eigenvalue is the discrete inf-sup constant β_{is}^h. □

Remark 3.160 (A Practical Aspect) The usual approach consists in using a subspace of $L^2(\Omega)$ for the discrete pressure and not of $L_0^2(\Omega)$, see Remark 4.70, and to fix the additive constant for the pressure in a different way. Applying this approach, the kernel of the matrix B is spanned by the constant vectors and the matrix has rank deficiency one, see Lemma 4.71. Consequently, the smallest eigenvalue of (3.150) is zero. In this case, the square root of the second smallest eigenvalue of (3.150) is the discrete inf-sup constant β_{is}^h. □

3.7 The Helmholtz Decomposition

Remark 3.161 (Contents) This section introduces a decomposition of vector fields in $L^2(\Omega)$ which is of importance for incompressible flow problems. With the help of this decomposition, one can describe, e.g., the impact of certain body forces on the velocity and pressure precisely, compare Remark 4.12. □

Definition 3.162 (Rotation and Curl) There are different operators which are called rotation or curl operator. Their definition depends on the dimension of the vector field.

Let $\boldsymbol{v}(\boldsymbol{x}) = (v_1(\boldsymbol{x}), v_2(\boldsymbol{x}), v_3(\boldsymbol{x}))^T$, $\boldsymbol{x} = (x, y, z)^T$, be a vector field in a domain $\Omega \subset \mathbb{R}^3$ which is sufficiently regular. Then the rotation or the curl of $\boldsymbol{v}(\boldsymbol{x})$ is defined by

$$\nabla \times \boldsymbol{v}(\boldsymbol{x}) = \det\begin{pmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \partial_x & \partial_y & \partial_z \\ v_1(\boldsymbol{x}) & v_2(\boldsymbol{x}) & v_3(\boldsymbol{x}) \end{pmatrix} = \begin{pmatrix} \partial_y v_3 - \partial_z v_2 \\ \partial_z v_1 - \partial_x v_3 \\ \partial_x v_2 - \partial_y v_1 \end{pmatrix}(\boldsymbol{x}). \tag{3.151}$$

A vector field $\boldsymbol{v}(\boldsymbol{x}) = (v_1(\boldsymbol{x}), v_2(\boldsymbol{x}))^T$, $\boldsymbol{x} = (x, y)^T$, in a two-dimensional domain Ω can be extended formally to a vector field with three values by $\boldsymbol{v}(\boldsymbol{x}) =$

$(v_1(\boldsymbol{x}), v_2(\boldsymbol{x}), 0)^T$. Then, the first two components in (3.151) vanish. The third component of the right-hand side in (3.151) is denoted by

$$\text{curl } \boldsymbol{v}(\boldsymbol{x}) = \left(\partial_x v_2 - \partial_y v_1\right)(\boldsymbol{x}). \tag{3.152}$$

The curl of a scalar field $v(\boldsymbol{x})$, $\boldsymbol{x} = (x, y)^T$, given in a two-dimensional domain, is defined by

$$\textbf{curl } v(\boldsymbol{x}) = \begin{pmatrix} -\partial_y v \\ \partial_x v \end{pmatrix}(\boldsymbol{x}). \tag{3.153}$$

□

Remark 3.163 (Properties of the Rotation or Curl Operator) Straightforward calculations show that it holds in two dimensions for a sufficiently smooth function

$$\text{curl } (\textbf{curl } v)(\boldsymbol{x}) = -\Delta v(\boldsymbol{x}). \tag{3.154}$$

In three dimensions, one has

$$\nabla \times (\nabla \times \boldsymbol{v})(\boldsymbol{x}) = -\Delta \boldsymbol{v}(\boldsymbol{x}) + \nabla (\nabla \cdot \boldsymbol{v})(\boldsymbol{x}). \tag{3.155}$$

Extending a two-dimensional vector field formally to a three-dimensional field, then also (3.155) is valid. Furthermore, on has

$$\text{curl } \nabla v(\boldsymbol{x}) = 0, \quad \nabla \times \nabla v(\boldsymbol{x}) = \boldsymbol{0} \tag{3.156}$$

and

$$\nabla \cdot (\textbf{curl } v)(\boldsymbol{x}) = 0, \quad \nabla \cdot (\nabla \times \boldsymbol{v})(\boldsymbol{x}) = 0. \tag{3.157}$$

Because of property (3.156), gradient fields are called irrotational. Other identities are given by

$$(\nabla \times \boldsymbol{v}) \times \boldsymbol{w} = (\boldsymbol{w} \cdot \nabla)\boldsymbol{v} - \nabla(\boldsymbol{v} \cdot \boldsymbol{w}) + (\nabla \boldsymbol{w})^T \boldsymbol{v} \tag{3.158}$$

and

$$\nabla \times (\boldsymbol{v} \times \boldsymbol{w}) = (\boldsymbol{w} \cdot \nabla)\boldsymbol{v} - (\boldsymbol{v} \cdot \nabla)\boldsymbol{w} + (\nabla \cdot \boldsymbol{w})\boldsymbol{v} - (\nabla \cdot \boldsymbol{v})\boldsymbol{w}. \tag{3.159}$$

□

Remark 3.164 (Spaces with the Curl Operator and Integration by Parts) Spaces that take into account curl operators are defined by

$$H(\mathbf{curl}\,, \Omega) = \left\{ v \in L^2(\Omega) \ : \ \mathbf{curl}\ v \in L^2(\Omega) \right\} \quad \text{if } d = 2,$$
$$H(\mathbf{curl}\,, \Omega) = \left\{ \boldsymbol{v} \in L^2(\Omega) \ : \ \nabla \times \boldsymbol{v} \in L^2(\Omega) \right\} \quad \text{if } d = 3.$$

If Ω is a bounded domain with Lipschitz boundary, then an integration by parts formula holds, see Girault and Raviart (1986, p. 34). Consider first the two-dimensional case. Let $\boldsymbol{v} \in H(\mathbf{curl}\,, \Omega)$, then it is

$$(\text{curl}\ \boldsymbol{v}, \phi) = (\boldsymbol{v}, \mathbf{curl}\ \phi) + \int_{\partial\Omega} (\boldsymbol{v} \cdot \boldsymbol{t}_1 \phi)(s)\, ds \quad \forall\ \phi \in H^1(\Omega), \tag{3.160}$$

where $\boldsymbol{t}_1(s)$ is the unit tangential vector at the boundary defined by $\boldsymbol{t}_1 = (-n_2, n_1)^T$, with $\boldsymbol{n} = (n_1, n_2)^T$ being the unit outer normal. In three dimensions, one has for all $\boldsymbol{v} \in H(\mathbf{curl}\,, \Omega)$

$$(\nabla \times \boldsymbol{v}, \boldsymbol{\phi}) = (\boldsymbol{v}, \nabla \times \boldsymbol{\phi}) + \int_{\partial\Omega} ((\boldsymbol{v} \times \boldsymbol{n}) \cdot \boldsymbol{\phi})(\boldsymbol{s})\, ds \quad \forall\ \boldsymbol{\phi} \in H^1(\Omega). \tag{3.161}$$

From (3.158), one obtains for $\boldsymbol{u} \in V_{\text{div}}$ and $\boldsymbol{v}, \boldsymbol{w}$ such that $\boldsymbol{v} \times \boldsymbol{w} \in L^2(\Omega)$, using integration by parts,

$$((\nabla \times \boldsymbol{v}) \times \boldsymbol{w}, \boldsymbol{u}) = ((\boldsymbol{w} \cdot \nabla)\, \boldsymbol{v}, \boldsymbol{u}) - ((\boldsymbol{u} \cdot \nabla)\, \boldsymbol{v}, \boldsymbol{w}).$$

□

Lemma 3.165 (Criterion for a Vector-valued Distribution to be the Gradient of a Function from $L^2(\Omega)$) *Let $\Omega \subset \mathbb{R}^d$, $d \geq 2$, be a bounded Lipschitz domain and let $\Omega_0 \subseteq \Omega$, $\Omega_0 \neq \emptyset$, be any subdomain. Suppose $\boldsymbol{v} \in W^{-1,2}(\Omega)$ satisfies*

$$\boldsymbol{v}(\psi) = 0 \quad \forall\ \psi \in C_{0,\text{div}}^\infty(\Omega),$$

where $C_{0,\text{div}}^\infty(\Omega)$ is defined in (A.7). Then there exists a unique $r \in L^2(\Omega)$ which satisfies

$$\int_{\Omega_0} r(\boldsymbol{x})\, d\boldsymbol{x} = 0 \quad \textit{and} \quad \boldsymbol{v} = \nabla r$$

in the sense of distributions. It holds

$$\|r\|_{L^2(\Omega)} \leq C_1 \|\boldsymbol{v}\|_{W^{-1,2}(\Omega)} \leq C_1 C_2 \|r\|_{L^2(\Omega)},$$

where $C_1 = C_1(\Omega_0, \Omega) > 0$ and $C_2 = C_2(d) > 0$.

Proof The proof can be found in Sohr (2001, p. 76). ■

Lemma 3.166 (Existence of a Solution of the Equation $\nabla \cdot \boldsymbol{v} = g$) *Let $\Omega \subset \mathbb{R}^d$, $d \geq 2$, be a bounded domain with Lipschitz boundary.*

For each $g \in W^{-1,2}(\Omega)$ there exists at least one $\boldsymbol{v} \in L^2(\Omega)$ with $\nabla \cdot \boldsymbol{v} = g$ in the sense of distributions and

$$\|\boldsymbol{v}\|_{L^2(\Omega)} \leq C \|g\|_{W^{-1,2}(\Omega)}$$

with $C = C(\Omega) > 0$.

For each $g \in L^2(\Omega)$ with $\int_\Omega g(\boldsymbol{x})\, d\boldsymbol{x} = 0$, there is at least one $\boldsymbol{v} \in L^2(\Omega)$ with $\nabla \cdot \boldsymbol{v} = g$ in the sense of distributions, $\boldsymbol{v} \cdot \boldsymbol{n}|_{\partial\Omega} = 0$ in the sense of (generalized) traces, and

$$\|\boldsymbol{v}\|_{L^2(\Omega)} \leq C \|g\|_{L^2(\Omega)} = C \|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)}$$

with $C = C(\Omega) > 0$.

Proof The proof can be found in Sohr (2001, p. 80). ∎

Remark 3.167 (On Lemma 3.166)

- The statements of Lemma 3.166 are generalizations of Corollary 3.44 in the sense that less regularity is assumed.
- In Sohr (2001), the results of the two previous lemmas are formulated for spaces $W^{-1,q}(\Omega)$ and $L^q(\Omega)$ with $q \in (1, \infty)$. Then, some of the constants in the estimates of Lemmas 3.165 and 3.166 depend on q.

□

Theorem 3.168 (Helmholtz Decomposition of a Vector Field in $L^2(\Omega)$) *Let $\Omega \subset \mathbb{R}^d$, $d \geq 2$, be a bounded Lipschitz domain. Then, each $\boldsymbol{v} \in L^2(\Omega)$ has a unique decomposition*

$$\boldsymbol{v} = \boldsymbol{w} + \nabla r, \tag{3.162}$$

with $\boldsymbol{w} \in H_{\text{div}}(\Omega)$ and $\nabla r \in G(\Omega)$, where the space $H_{\text{div}}(\Omega)$ is defined in (3.39) and

$$G(\Omega) = \left\{\boldsymbol{v} \in L^2(\Omega) \ : \ \exists r \in L^2(\Omega) \ : \ \boldsymbol{v} = \nabla r\right\}.$$

The spaces $H_{\text{div}}(\Omega)$ and $G(\Omega)$ are orthogonal in $L^2(\Omega)$, i.e.,

$$G(\Omega) = H_{\text{div}}(\Omega)^\perp.$$

Consequently, it is $(\boldsymbol{w}, \nabla r) = 0$ and it holds

$$\|\boldsymbol{v}\|_{L^2(\Omega)}^2 = \|\boldsymbol{w}\|_{L^2(\Omega)}^2 + \|\nabla r\|_{L^2(\Omega)}^2. \tag{3.163}$$

Proof For the proof it is referred to Sohr (2001, p. 82). ∎

Remark 3.169 (Arbitrary Domains) The Helmholtz decomposition exists also for arbitrary domains, see Sohr (2001, Lemma 2.5.1), however the characterization of the space G changes, it is less special. □

Definition 3.170 (Helmholtz Projection) Using the Helmholtz decomposition (3.162), the Helmholtz projection is defined by

$$P_{\text{helm}} : L^2(\Omega) \to H_{\text{div}}(\Omega), \quad \boldsymbol{v} \mapsto \boldsymbol{w}.$$

□

Lemma 3.171 (Properties of the Helmholtz Projection) *The Helmholtz projection is a uniquely determined, bounded linear operator with* $\|P_{\text{helm}}\| \le 1$*, i.e.,*

$$\|P_{\text{helm}}\boldsymbol{v}\|_{L^2(\Omega)} \le \|\boldsymbol{v}\|_{L^2(\Omega)} \quad \forall\ \boldsymbol{v} \in L^2(\Omega). \tag{3.164}$$

It has the following properties

$$\begin{aligned} P_{\text{helm}}(\nabla r) &= 0, & (I - P_{\text{helm}})\,\boldsymbol{v} &= \nabla r, \\ P_{\text{helm}}^2\boldsymbol{v} &= P_{\text{helm}}\boldsymbol{v}, & (I - P_{\text{helm}})^2\,\boldsymbol{v} &= (I - P_{\text{helm}})\,\boldsymbol{v}, \end{aligned}$$

for all $\boldsymbol{v} \in L^2(\Omega)$*. Furthermore, the operator* P_{helm} *is selfadjoint, i.e.,*

$$(P_{\text{helm}}\boldsymbol{v}, \boldsymbol{g}) = (\boldsymbol{v}, P_{\text{helm}}\boldsymbol{g}) \quad \forall\ \boldsymbol{v}, \boldsymbol{g} \in L^2(\Omega).$$

Proof By Hilbert space theory, the projection operator P_{helm} is uniquely determined. The boundedness (3.164) follows directly from (3.163)

$$\|\boldsymbol{v}\|^2_{L^2(\Omega)} \ge \|\boldsymbol{w}\|^2_{L^2(\Omega)} = \|P_{\text{helm}}\boldsymbol{v}\|^2_{L^2(\Omega)}.$$

Then, the next four properties follow from (3.162) and the uniqueness of the Helmholtz decomposition. Finally, the last property follows from the orthogonality of $H_{\text{div}}(\Omega)$ and $G(\Omega)$. Let $\boldsymbol{v} = \boldsymbol{w} + \nabla r$ and $\boldsymbol{g} = \boldsymbol{w_g} + \nabla r_{\boldsymbol{g}}$ with $\boldsymbol{w}, \boldsymbol{w_g} \in H_{\text{div}}(\Omega)$, $r, r_{\boldsymbol{g}} \in G(\Omega)$, be the Helmholtz decompositions of $\boldsymbol{v}$ and $\boldsymbol{g}$, respectively. Then it follows that

$$\begin{aligned} (P_{\text{helm}}\boldsymbol{v}, \boldsymbol{g}) &= \left(\boldsymbol{w}, \boldsymbol{w_g} + \nabla r_{\boldsymbol{g}}\right) = \left(\boldsymbol{w}, \boldsymbol{w_g}\right) + \left(\boldsymbol{w}, \nabla r_{\boldsymbol{g}}\right) = \left(\boldsymbol{w}, \boldsymbol{w_g}\right) \\ &= \left(\boldsymbol{w} + \nabla r, \boldsymbol{w_g}\right) = (\boldsymbol{v}, P_{\text{helm}}\boldsymbol{g}). \end{aligned}$$

■

Remark 3.172 (Computation of the Helmholtz Decomposition) So far, the existence and uniqueness of the Helmholtz decomposition is known from Theorem 3.168. The next question is how the decomposition can be computed. It will turn out that under certain conditions on the domain Ω, the functions $\boldsymbol{w}(\boldsymbol{x})$ and $r(\boldsymbol{x})$ can be characterized as the solution of boundary value problems. □

Lemma 3.173 (Definition of r) *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be a bounded Lipschitz domain with boundary $\Gamma = \cup_{i=0,\ldots,k}\Gamma_i$, where Γ_0 is the exterior boundary of Ω and $\Gamma_1, \ldots, \Gamma_k$ are other components. Then, $r \in H^1(\Omega)/\mathbb{R}$ is the unique solution of*

$$(\nabla r, \nabla\phi) = (\boldsymbol{v}, \nabla\phi) \quad \forall\, \phi \in H^1(\Omega),$$

where

$$H^1(\Omega)/\mathbb{R} = \left\{ v \in H^1(\Omega) \ : \ \int_\Omega v(\boldsymbol{x})\, d\boldsymbol{x} = 0 \right\}.$$

If $\boldsymbol{v} \in H(\mathrm{div}, \Omega)$, then this problem can be interpreted as the following Poisson problem with Neumann boundary conditions

$$\begin{aligned} -\Delta r &= -\nabla \cdot \boldsymbol{v} \ \ \text{in } \Omega, \\ \nabla r \cdot \boldsymbol{n} &= \boldsymbol{v} \cdot \boldsymbol{n} \quad \text{on } \Gamma. \end{aligned} \tag{3.165}$$

Proof The proof can be found in Girault and Raviart (1986, p. 41). ■

Lemma 3.174 (Definition of $\boldsymbol{w}$ in Two Dimensions) *Let $\Omega \subset \mathbb{R}^2$ be a bounded, simply connected domain with Lipschitz boundary. Then it is*

$$\boldsymbol{w} = \mathbf{curl}\ \phi,$$

where $\phi \in V$ is the unique solution of

$$(\mathbf{curl}\ \phi, \mathbf{curl}\ \psi) = (\boldsymbol{v} - \nabla r, \mathbf{curl}\ \psi) \quad \forall\, \psi \in V. \tag{3.166}$$

The function $\phi(\boldsymbol{x})$ is called stream function. This problem has the following interpretation

$$\begin{aligned} -\Delta\phi &= \mathrm{curl}\ \boldsymbol{v} \ \ \text{in } \Omega, \\ \phi &= 0 \qquad \text{on } \Gamma, \end{aligned} \tag{3.167}$$

where the equation has to be understood in $H^{-1}(\Omega)$.

Proof Also the proof of this statement can be found in Girault and Raviart (1986, p. 40). Here, only the correspondence of the variational and the formulation (3.167) will be illustrated. Considering first the left-hand side of (3.166), then the integration by parts formula (3.160) and (3.154) yields

$$(\mathbf{curl}\ \phi, \mathbf{curl}\ \psi) = (\mathrm{curl}\ \mathbf{curl}\ \phi, \psi) + \int_{\partial\Omega} ((\mathbf{curl}\ \phi \cdot \boldsymbol{t}_1)\, \psi)\,(s)\, ds = (-\Delta\phi, \psi)$$

for all $\psi \in V$. The boundary integral vanishes since the trace of $\psi(\boldsymbol{x})$ is zero. Similarly, one gets for the right-hand side of (3.166), using (3.156),

$$(\boldsymbol{v} - \nabla r, \mathbf{curl}\ \psi) = (\text{curl}\ \boldsymbol{v} - \text{curl}\ \nabla r, \psi) = (\text{curl}\ \boldsymbol{v}, \psi)$$

for all $\psi \in V$. Hence, the formulation (3.167) follows. ■

Remark 3.175 (Stream Function) Let $\boldsymbol{v} \in H(\text{div}, \Omega)$ and let $\boldsymbol{v}(\boldsymbol{x})$ be weakly divergence-free. Then, integration by parts gives

$$0 = \int_\Omega \nabla \cdot \boldsymbol{v}(\boldsymbol{x})\, d\boldsymbol{x} = \int_\Gamma (\boldsymbol{v} \cdot \boldsymbol{n})(\boldsymbol{s})\, ds.$$

That means, for a simply connected Lipschitz domain, problem (3.165) has a homogeneous right-hand side. It is known that this Neumann problem has a unique solution in the quotient space $H^1(\Omega)/\mathbb{R}$, see Girault and Raviart (1986, p. 14). Obviously, the solution is $r(\boldsymbol{x}) = 0$ and hence it is in two dimensions

$$\boldsymbol{v} = \boldsymbol{w} = \mathbf{curl}\ \phi. \tag{3.168}$$

Given a weakly divergence-free vector field in two dimensions, the stream function can be computed with (3.168), which is often used for visualization purposes. □

Lemma 3.176 (Definition of $\boldsymbol{w}$ in Three Dimensions) *Let $\Omega \subset \mathbb{R}^3$ be a bounded, simply connected domain with Lipschitz boundary. Then*

$$\boldsymbol{w} = \nabla \times \boldsymbol{\phi},$$

where $\boldsymbol{\phi} \in H = \{\nabla \times \boldsymbol{\phi} : \boldsymbol{\phi} \in V\}$ is the unique solution of the boundary value problem

$$\begin{aligned} -\Delta \boldsymbol{\phi} &= \nabla \times \boldsymbol{v} \quad in\ \Omega, \\ (\nabla \times \boldsymbol{\phi} - \boldsymbol{v}) \cdot \boldsymbol{n} &= 0 \qquad on\ \Gamma, \end{aligned}$$

where the equation has to be understood in the sense of $H^{-1}(\Omega)$. The function $\boldsymbol{\phi}$ is called vector potential.

Proof For the proof, see Girault and Raviart (1986, p. 48). ■

Remark 3.177 (Multiple Connected Domains) The statement of Lemma 3.174 can be extended to domains with the properties from the assumptions of Lemma 3.173 and with the boundary conditions

$$\int_{\Gamma_i} (\boldsymbol{v} \cdot \boldsymbol{n})(\boldsymbol{s})\, d\boldsymbol{s} = 0, \quad i = 0, \ldots, k.$$

However, then the ansatz and test space are somewhat more complicated as well as the boundary condition of the corresponding problem, see Girault and Raviart (1986, p. 40). The definition of a stream function is still possible.

The statement of Lemma 3.176 is not true for multiple connected domains. There, $\boldsymbol{w}$ is not necessarily a curl vector. It may be the sum of a curl vector and some special vector which belongs to a finite-dimensional space, see Temam (1984, Appendix I). □

Remark 3.178 (Summary for Simply Connected Lipschitz Domains) For simply connected Lipschitz domains in $\mathbb{R}^d$, $d \in \{2, 3\}$, every vector field in $L^2(\Omega)$ is the sum of the curl of a stream function (in two dimension) or vector potential (in three dimensions) and a gradient. □

Lemma 3.179 (Estimating the Divergence by the Gradient for Functions from $H_0^1(\Omega)$) *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, and let $\boldsymbol{v} \in H_0^1(\Omega)$, then it holds*

$$\|\nabla \boldsymbol{v}\|_{L^2(\Omega)}^2 = \|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)}^2 + \|\nabla \times \boldsymbol{v}\|_{L^2(\Omega)}^2 \tag{3.169}$$

and consequently

$$\|\nabla \cdot \boldsymbol{v}\|_{L^2(\Omega)} \le \|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \quad \forall\ \boldsymbol{v} \in H_0^1(\Omega). \tag{3.170}$$

Proof The proof is based on the identity (3.155). Considering $\boldsymbol{v} \in H_0^1(\Omega)$, this identity can be transformed into a weak form by multiplication with a test function $\boldsymbol{w} \in H_0^1(\Omega)$ and applying integration by parts

$$(\nabla \boldsymbol{v}, \nabla \boldsymbol{w}) = (\nabla \cdot \boldsymbol{v}, \nabla \cdot \boldsymbol{w}) + (\nabla \times \boldsymbol{v}, \nabla \times \boldsymbol{w}) \quad \forall\ \boldsymbol{w} \in H_0^1(\Omega). \tag{3.171}$$

The derivation of the first two terms is standard. For deriving the last term, the integration by parts formula (3.161) can be applied. However, the derivation can be also checked with a straightforward calculation. Using (3.151) gives

$$\nabla \times \nabla \times \boldsymbol{v} = \begin{pmatrix} \partial_y\left(\partial_x v_2 - \partial_y v_1\right) - \partial_z\left(\partial_z v_1 - \partial_x v_3\right) \\ \partial_z\left(\partial_y v_3 - \partial_z v_2\right) - \partial_x\left(\partial_x v_2 - \partial_y v_1\right) \\ \partial_x\left(\partial_z v_1 - \partial_x v_3\right) - \partial_y\left(\partial_y v_3 - \partial_z v_2\right) \end{pmatrix}.$$

Applying integration by parts, boundary integrals will not appear for test functions $\boldsymbol{w} \in H_0^1(\Omega)$. One obtains, considering the individual terms,

$$\begin{aligned}(\nabla \times \nabla \times \boldsymbol{v}, \boldsymbol{w}) &= -\left(\partial_x v_2 - \partial_y v_1, \partial_y w_1\right) + (\partial_z v_1 - \partial_x v_3, \partial_z w_1) - \left(\partial_y v_3 - \partial_z v_2, \partial_z w_2\right) \\ &\quad + \left(\partial_x v_2 - \partial_y v_1, \partial_x w_2\right) - (\partial_z v_1 - \partial_x v_3, \partial_x w_3) + \left(\partial_y v_3 - \partial_z v_2, \partial_y w_3\right) \\ &= \left(\partial_y v_3 - \partial_z v_2, \partial_y w_3 - \partial_z w_2\right) + (\partial_z v_1 - \partial_x v_3, \partial_z w_1 - \partial_x w_3) \\ &\quad + \left(\partial_x v_2 - \partial_y v_1, \partial_x w_2 - \partial_y w_1\right) \\ &= (\nabla \times \boldsymbol{v}, \nabla \times \boldsymbol{w}) \,.\end{aligned}$$

Inserting now $\boldsymbol{w} = \boldsymbol{v} \in H_0^1(\Omega)$ in (3.171) gives (3.169). ■

Chapter 4
The Stokes Equations

Remark 4.1 (Motivation) The Stokes equations model the simplest incompressible flow problems. These problems are steady-state and the convective term can be neglected. Hence, the arising model is linear. Thus, the only difficulty which remains from the problems mentioned in Remark 2.19 is the coupling of velocity and pressure.

The analysis of the Stokes equations and of finite element discretizations of these equations introduces already important techniques which will be used also in the analysis of more complicated problems. □

4.1 The Continuous Equations

Remark 4.2 (The Stokes Equations) Consider a stationary flow, i.e., $\partial_t \boldsymbol{u} = \boldsymbol{0}$. If the flow is in addition very slow, i.e., the Reynolds number is very small, then the viscous term $Re^{-1}\Delta\boldsymbol{u}$ dominates the convective term $(\boldsymbol{u}\cdot\nabla)\boldsymbol{u}$ and the convective term can be neglected. The resulting momentum equation can be scaled by the Reynolds number, defining a new pressure and right-hand side in this way. One obtains the so-called Stokes equations

$$\begin{aligned} -\Delta\boldsymbol{u} + \nabla p &= \boldsymbol{f} \text{ in } \Omega, \\ \nabla\cdot\boldsymbol{u} &= 0 \text{ in } \Omega, \end{aligned} \tag{4.1}$$

that has to be equipped with appropriate boundary conditions.

The theory from Sects. 3.1 and 3.2 will be applied here to study system (4.1). For simplicity of presentation, the Stokes equations will be considered with homogeneous Dirichlet boundary conditions $\boldsymbol{u} = \boldsymbol{0}$ on Γ. □

Remark 4.3 (The Weak Form of Stokes Equations) The weak form of the Stokes equations equipped with homogeneous Dirichlet boundary conditions is obtained

V. John, *Finite Element Methods for Incompressible Flow Problems*, Springer Series in Computational Mathematics 51, DOI 10.1007/978-3-319-45750-5_4

in the usual way by multiplying the equations with test functions, integrating these equations on Ω, and applying integration by parts to transfer derivatives from the solution to the test functions. One obtains the following problem: Given $\boldsymbol{f} \in H^{-1}(\Omega)$, find $(\boldsymbol{u}, p) \in H_0^1(\Omega) \times L_0^2(\Omega)$ such that

$$\begin{aligned} (\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, p) &= \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{H^{-1}(\Omega), H_0^1(\Omega)} \quad \forall\ \boldsymbol{v} \in H_0^1(\Omega), \\ -(\nabla \cdot \boldsymbol{u}, q) &= 0 \qquad\qquad\qquad\qquad \forall\ q \in L_0^2(\Omega). \end{aligned} \tag{4.2}$$

This weak form can be cast into the framework of the abstract linear saddle point problem by setting $V = H_0^1(\Omega)$ and $Q = L_0^2(\Omega)$ in (3.4), equipped with the norms $\|\cdot\|_V = |\cdot|_{H^1(\Omega)}$ and $\|\cdot\|_Q = \|\cdot\|_{L^2(\Omega)}$,

$$a(\boldsymbol{u}, \boldsymbol{v}) = (\nabla \boldsymbol{u}, \nabla \boldsymbol{v}), \quad b(\boldsymbol{v}, q) = -(\nabla \cdot \boldsymbol{v}, q), \quad r = 0.$$

An equivalent formulation of (4.2) is as follows: Find $(\boldsymbol{u}, p) \in V \times Q$ such that

$$a(\boldsymbol{u}, \boldsymbol{v}) + b(\boldsymbol{v}, p) - b(\boldsymbol{u}, q) = \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{V', V} \quad \forall\ (\boldsymbol{v}, q) \in V \times Q. \tag{4.3}$$

If (4.2) holds, one gets (4.3) by subtracting the second equation from the first equation in (4.2). In turn, if (4.3) is valid, then the individual equations of (4.2) are obtained by considering in (4.3) the sets $\{(\boldsymbol{v}, 0)\}$ and $\{(\boldsymbol{0}, q)\}$ as test functions.

Let $V_0 = V_{\text{div}}$, the space of weakly divergence-free functions defined in (3.38). Then, the associated problem to (4.2), which corresponds to (3.12), is: Find $\boldsymbol{u} \in V_{\text{div}}$ such that

$$(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) = \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{V', V} \quad \forall\ \boldsymbol{v} \in V_{\text{div}}. \tag{4.4}$$

□

Remark 4.4 (The Deformation Tensor Form of the Viscous Term) Starting point for a variational formulation of the viscous term might be also the term $-2\nabla \cdot (\mathbb{D}(\boldsymbol{u}))$ which is equivalent to $-\Delta \boldsymbol{u}$, see Remark 2.20. Multiplying this term with a test function $\boldsymbol{v} \in V$, applying integration by parts, utilizing that the $L^2(\Omega)$ inner product of tensors is defined component by component, and using that $\mathbb{D}(\boldsymbol{u})$ is symmetric yields

$$\begin{aligned} -2\left(\nabla \cdot (\mathbb{D}(\boldsymbol{u})), \boldsymbol{v}\right) &= 2\left(\mathbb{D}(\boldsymbol{u}), \nabla \boldsymbol{v}\right) = \left(\mathbb{D}(\boldsymbol{u}), \nabla \boldsymbol{v}\right) + \left(\mathbb{D}(\boldsymbol{u}), \nabla \boldsymbol{v}\right) \\ &= \left(\mathbb{D}(\boldsymbol{u}), \nabla \boldsymbol{v}\right) + \left((\mathbb{D}(\boldsymbol{u}))^T, (\nabla \boldsymbol{v})^T\right) \\ &= \left(\mathbb{D}(\boldsymbol{u}), \nabla \boldsymbol{v}\right) + \left(\mathbb{D}(\boldsymbol{u}), (\nabla \boldsymbol{v})^T\right) \\ &= 2\left(\mathbb{D}(\boldsymbol{u}), \mathbb{D}(\boldsymbol{v})\right). \end{aligned} \tag{4.5}$$

□

Lemma 4.5 (The Norm of the Bilinear Form $a(\cdot,\cdot)$) *For the bilinear form $a(\cdot,\cdot)$ associated with the Stokes problem it holds $\|a\| = 1$.*

Proof One gets with the Cauchy–Schwarz inequality (A.10)

$$\|a\| = \sup_{\boldsymbol{v},\boldsymbol{w}\in V\setminus\{\boldsymbol{0}\}} \frac{(\nabla \boldsymbol{v}, \nabla \boldsymbol{w})}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \|\nabla \boldsymbol{w}\|_{L^2(\Omega)}} \le \sup_{\boldsymbol{v},\boldsymbol{w}\in V\setminus\{\boldsymbol{0}\}} \frac{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \|\nabla \boldsymbol{w}\|_{L^2(\Omega)}}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \|\nabla \boldsymbol{w}\|_{L^2(\Omega)}} = 1.$$

Choosing $\boldsymbol{v} = \boldsymbol{w}$ shows that the supremum is not smaller than 1. ■

Theorem 4.6 (Existence and Uniqueness of a Solution of the Stokes Equations) *Let Ω be a bounded domain in $\mathbb{R}^d$, $d \in \{2,3\}$, with a Lipschitz continuous boundary Γ and let $\boldsymbol{f} \in H^{-1}(\Omega)$. Then, there exists a unique pair $(\boldsymbol{u}, p) \in H_0^1(\Omega) \times L_0^2(\Omega)$ that solves (4.2).*

Proof The bilinear form $b(\cdot,\cdot)$ satisfies the inf-sup condition, see Theorem 3.46. In addition, the bilinear form $a(\cdot,\cdot)$ is V_{div}-elliptic since

$$a(\boldsymbol{v}, \boldsymbol{v}) = |\boldsymbol{v}|_{H^1(\Omega)}^2 = \|\boldsymbol{v}\|_V^2 \quad \forall\ \boldsymbol{v} \in V \supset V_{\mathrm{div}}. \tag{4.6}$$

The statement of the theorem follows now from Lemma 3.19. ■

Lemma 4.7 (Stability of the Solution) *Let the conditions of Theorem 4.6 be given. Then, the solution $(\boldsymbol{u}, p)$ of the Stokes problem (4.2) depends continuously on the right-hand side*

$$\|\boldsymbol{u}\|_V = \|\nabla \boldsymbol{u}\|_{L^2(\Omega)} \le \|\boldsymbol{f}\|_{H^{-1}(\Omega)}, \tag{4.7}$$

$$\|p\|_Q = \|p\|_{L^2(\Omega)} \le \frac{2}{\beta_{\mathrm{is}}} \|\boldsymbol{f}\|_{H^{-1}(\Omega)}. \tag{4.8}$$

If $\boldsymbol{f} \in L^2(\Omega)$, then it holds

$$\|\boldsymbol{u}\|_V \le C \|P_{\mathrm{helm}} \boldsymbol{f}\|_{L^2(\Omega)}, \tag{4.9}$$

where the constant comes from the Poincaré inequality (A.12).

Proof Inserting $\boldsymbol{u}$ as test function in (4.4) gives

$$(\nabla \boldsymbol{u}, \nabla \boldsymbol{u}) = \langle \boldsymbol{f}, \boldsymbol{u}\rangle_{V',V}.$$

Applying the dual estimate yields

$$\|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^2 \le \|\boldsymbol{f}\|_{H^{-1}(\Omega)} \|\nabla \boldsymbol{u}\|_{L^2(\Omega)}.$$

In the case $\|\nabla \boldsymbol{u}\|_{L^2(\Omega)} = 0$, the estimate for the velocity is trivially true. Otherwise, division with $\|\nabla \boldsymbol{u}\|_{L^2(\Omega)}$ leads to (4.7).

For the estimate of the pressure, the inf-sup condition (3.48), the equation (4.2), the estimate of the dual pairing, and the Cauchy–Schwarz inequality (A.10) are utilized

$$
\begin{aligned}
\|p\|_{L^2(\Omega)} &\le \frac{1}{\beta_{\mathrm{is}}} \sup_{\boldsymbol{v}\in V\setminus\{\boldsymbol{0}\}} \frac{-(\nabla\cdot\boldsymbol{v},p)}{\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}} \\
&= \frac{1}{\beta_{\mathrm{is}}} \sup_{\boldsymbol{v}\in V\setminus\{\boldsymbol{0}\}} \frac{\langle \boldsymbol{f},\boldsymbol{v}\rangle_{V',V} - (\nabla\boldsymbol{u},\nabla\boldsymbol{v})}{\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}} \\
&\le \frac{1}{\beta_{\mathrm{is}}} \sup_{\boldsymbol{v}\in V\setminus\{\boldsymbol{0}\}} \frac{\|\boldsymbol{f}\|_{H^{-1}(\Omega)} \|\nabla\boldsymbol{v}\|_{L^2(\Omega)} + \|\nabla\boldsymbol{u}\|_{L^2(\Omega)} \|\nabla\boldsymbol{v}\|_{L^2(\Omega)}}{\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}} \\
&= \frac{1}{\beta_{\mathrm{is}}} \left(\|\boldsymbol{f}\|_{H^{-1}(\Omega)} + \|\nabla\boldsymbol{u}\|_{L^2(\Omega)}\right).
\end{aligned}
$$

Inserting now the stability estimate (4.7) for the velocity gives the estimate (4.8) for the pressure.

Let $\boldsymbol{f} \in L^2(\Omega)$, then the application of Helmholtz decomposition (3.162)

$$
\boldsymbol{f} = P_{\mathrm{helm}}\boldsymbol{f} + \nabla r,
$$

using the Helmholtz projection from Definition 3.170, integration by parts, $\boldsymbol{u} \in V_{\mathrm{div}}$, the Cauchy–Schwarz inequality, and Poincaré inequality yields

$$
\begin{aligned}
\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^2 &= (P_{\mathrm{helm}}\boldsymbol{f},\boldsymbol{u}) + (\nabla r,\boldsymbol{u}) = (P_{\mathrm{helm}}\boldsymbol{f},\boldsymbol{u}) \\
&\le \|P_{\mathrm{helm}}\boldsymbol{f}\|_{L^2(\Omega)} \|\boldsymbol{u}\|_{L^2(\Omega)} \le C\,\|P_{\mathrm{helm}}\boldsymbol{f}\|_{L^2(\Omega)} \|\nabla\boldsymbol{u}\|_{L^2(\Omega)}.
\end{aligned}
$$

Division by $\|\nabla \boldsymbol{u}\|_{L^2(\Omega)}$ proves the last statement (4.9) of the lemma. ∎

Remark 4.8 (On the Implication of the Inf-Sup Condition) It can be shown with a straightforward calculation that the inf-sup condition guarantees the uniqueness of the pressure.

Since the weak form of the homogeneous Stokes equations

$$
\begin{aligned}
a(\boldsymbol{u},\boldsymbol{v}) + b(\boldsymbol{v},p) &= 0 \quad \forall\, \boldsymbol{v} \in V, \\
b(\boldsymbol{u},q) &= 0 \quad \forall\, q \in Q,
\end{aligned}
\tag{4.10}
$$

is a linear system, one has to show for uniqueness of a solution that $\boldsymbol{f} = \boldsymbol{0}$ implies $\boldsymbol{u} = \boldsymbol{0}$ and $p = 0$. Assume there is a solution $(\boldsymbol{u},p) \in V \times Q$ of (4.10). Taking $(\boldsymbol{u},p)$ as test functions gives

$$
\begin{aligned}
a(\boldsymbol{u},\boldsymbol{u}) + b(\boldsymbol{u},p) &= 0, \\
b(\boldsymbol{u},p) &= 0.
\end{aligned}
$$

Including the second equation in the first one leads to $a(\boldsymbol{u}, \boldsymbol{u}) = 0$. The ellipticity of $a(\cdot, \cdot)$ in V, see (4.6), implies $\boldsymbol{u} = \boldsymbol{0}$. Up to this point, the inf-sup condition was not needed.

Considering now the first equation of (4.10) with $\boldsymbol{u} = \boldsymbol{0}$ gives

$$b(\boldsymbol{v}, p) = 0 \quad \forall\ \boldsymbol{v} \in V.$$

It follows that

$$\sup_{\boldsymbol{v} \in V \setminus \{\boldsymbol{0}\}} \frac{b(\boldsymbol{v}, p)}{\|\boldsymbol{v}\|_V} = 0.$$

The inf-sup condition written in the form (3.47) implies now immediately $p = 0$ in $L_0^2(\Omega)$ since $\beta_{\mathrm{is}} > 0$. □

Remark 4.9 (Non-homogeneous Dirichlet Boundary Conditions) Consider the Stokes equations with the boundary conditions

$$\boldsymbol{u} = \boldsymbol{g} \text{ on } \Gamma,$$

with $\boldsymbol{g} \in H^{1/2}(\Gamma)$ such that the compatibility condition (2.33) is satisfied.

The existence and uniqueness of a solution of the Stokes equations with non-homogeneous Dirichlet boundary condition can be proved also with Theorem 4.6. From the Trace Theorem, Theorem A.34, it follows that there is a function $\boldsymbol{u}_0 \in H^1(\Omega)$ with $\boldsymbol{u}_0 = \boldsymbol{g}$ on Γ. The function $\boldsymbol{u} - \boldsymbol{u}_0$ fulfills the Stokes equations with homogeneous Dirichlet boundary conditions and a modified right-hand side. The right-hand side of the momentum equation is still in $H^{-1}(\Omega)$ and the right-hand side of the mass balance might be inhomogeneous. Note that the theory of Sects. 3.1 and 3.2 does not require the homogeneity of the right-hand side of the second equation of the saddle point problem. Theorem 4.6 shows the existence and uniqueness of a solution to this Stokes equations with new right-hand side. Adding $\boldsymbol{u}_0$ to this solution gives the existence of a solution to the Stokes problem with inhomogeneous Dirichlet boundary conditions.

The uniqueness is proved in an indirect way by assuming the existence of two different solutions. The difference of these solutions solves a Stokes problem with homogeneous right-hand side and homogeneous Dirichlet boundary conditions. Theorem 4.6 shows that the solution of this problem is $(\boldsymbol{0}, 0)$. Hence, the existence of two different solutions is not possible. □

Remark 4.10 (Other Boundary Conditions) Investigations of the well-posedness of the Stokes problem with others than Dirichlet boundary conditions can be found, e.g., in Ern and Guermond (2004, §4.1.4). □

Lemma 4.11 (Babuška Inf-Sup Condition for the Stokes Problem) *It holds*

$$\inf_{\substack{(\boldsymbol{w},r)\in(V\times Q)\\(\boldsymbol{w},r)\neq(\boldsymbol{0},0)}}\sup_{\substack{(\boldsymbol{v},q)\in(V\times Q)\\(\boldsymbol{v},q)\neq(\boldsymbol{0},0)}}\frac{(\nabla\boldsymbol{w},\nabla\boldsymbol{v})-(\nabla\cdot\boldsymbol{v},r)+(\nabla\cdot\boldsymbol{w},q)}{\left(\|\nabla\boldsymbol{w}\|_{L^2(\Omega)}+\|r\|_{L^2(\Omega)}\right)\left(\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}+\|q\|_{L^2(\Omega)}\right)} \geq \beta_{\mathrm{is,Bab}} > 0 \tag{4.11}$$

with

$$\frac{1}{\beta_{\mathrm{is,Bab}}} = \left(1+\frac{2}{\beta_{\mathrm{is}}}\right)+\frac{1}{\beta_{\mathrm{is}}}\left(1+\frac{1}{\beta_{\mathrm{is}}}\right).$$

Proof Let $(\boldsymbol{w},r) \in V\times Q$ be arbitrary. Inserting this pair in the left-hand side operator of the Stokes equations gives

$$(\nabla\boldsymbol{w},\nabla\boldsymbol{v})-(\nabla\cdot\boldsymbol{v},r)=\langle\boldsymbol{f},\boldsymbol{v}\rangle_{V',V},\quad (\nabla\cdot\boldsymbol{w},q)=(g,q)\quad\forall\ (\boldsymbol{v},q)\in V\times Q \tag{4.12}$$

for some pair $(\boldsymbol{f},g)\in V'\times Q$. Let $\boldsymbol{w}_{\mathrm{div}}\in V_{\mathrm{div}}$ be the orthogonal projection of $\boldsymbol{w}$ with respect to the inner product of V. Hence, $(\boldsymbol{w}-\boldsymbol{w}_{\mathrm{div}})\in(V_{\mathrm{div}})^{\perp}$. From Corollary 3.44, it follows that $\nabla\cdot(\boldsymbol{w}-\boldsymbol{w}_{\mathrm{div}})\in Q$ and one gets from (4.12), $\nabla\cdot\boldsymbol{w}_{\mathrm{div}}=0$, and the Cauchy–Schwarz inequality (A.10)

$$\begin{aligned}(\nabla\cdot\boldsymbol{w},\nabla\cdot(\boldsymbol{w}-\boldsymbol{w}_{\mathrm{div}})) &= (\nabla\cdot(\boldsymbol{w}-\boldsymbol{w}_{\mathrm{div}}),\nabla\cdot(\boldsymbol{w}-\boldsymbol{w}_{\mathrm{div}}))\\ &= \|\nabla\cdot(\boldsymbol{w}-\boldsymbol{w}_{\mathrm{div}})\|_{L^2(\Omega)}^2\\ &= (g,\nabla\cdot(\boldsymbol{w}-\boldsymbol{w}_{\mathrm{div}}))\\ &\leq \|g\|_{L^2(\Omega)}\|\nabla\cdot(\boldsymbol{w}-\boldsymbol{w}_{\mathrm{div}})\|_{L^2(\Omega)}.\end{aligned}$$

Using this estimate and (3.49), one obtains

$$\|\nabla(\boldsymbol{w}-\boldsymbol{w}_{\mathrm{div}})\|_{L^2(\Omega)}\leq\frac{1}{\beta_{\mathrm{is}}}\|\nabla\cdot(\boldsymbol{w}-\boldsymbol{w}_{\mathrm{div}})\|_{L^2(\Omega)}\leq\frac{1}{\beta_{\mathrm{is}}}\|g\|_{L^2(\Omega)}.$$

Choosing $\boldsymbol{v}=\boldsymbol{w}_{\mathrm{div}}$ in (4.12) gives

$$(\nabla\boldsymbol{w},\nabla\boldsymbol{w}_{\mathrm{div}})-(\nabla\cdot\boldsymbol{w}_{\mathrm{div}},r)=(\nabla\boldsymbol{w},\nabla\boldsymbol{w}_{\mathrm{div}})=\langle\boldsymbol{f},\boldsymbol{w}_{\mathrm{div}}\rangle_{V',V}.$$

From the orthogonality, it follows that

$$(\nabla\boldsymbol{w},\nabla\boldsymbol{w}_{\mathrm{div}})=(\nabla(\boldsymbol{w}-\boldsymbol{w}_{\mathrm{div}}),\nabla\boldsymbol{w}_{\mathrm{div}})+(\nabla\boldsymbol{w}_{\mathrm{div}},\nabla\boldsymbol{w}_{\mathrm{div}})=(\nabla\boldsymbol{w}_{\mathrm{div}},\nabla\boldsymbol{w}_{\mathrm{div}})$$

such that one obtains with the estimate for the dual pairing

$$\|\nabla\boldsymbol{w}_{\mathrm{div}}\|_{L^2(\Omega)}\leq\|\boldsymbol{f}\|_{H^{-1}(\Omega)}.$$

Now, one gets with the triangle inequality

$$\|\nabla \boldsymbol{w}\|_{L^2(\Omega)} \le \|\nabla (\boldsymbol{w} - \boldsymbol{w}_{\text{div}})\|_{L^2(\Omega)} + \|\nabla \boldsymbol{w}_{\text{div}}\|_{L^2(\Omega)} \le \frac{1}{\beta_{\text{is}}} \|g\|_{L^2(\Omega)} + \|\boldsymbol{f}\|_{H^{-1}(\Omega)}. \tag{4.13}$$

Using the same techniques like for the stability estimate for the pressure in Lemma 4.7 yields, together with (4.13),

$$\|r\|_{L^2(\Omega)} \le \frac{1}{\beta_{\text{is}}} \left(\|\boldsymbol{f}\|_{H^{-1}(\Omega)} + \|\nabla \boldsymbol{w}\|_{L^2(\Omega)}\right) \le \frac{2}{\beta_{\text{is}}} \|\boldsymbol{f}\|_{H^{-1}(\Omega)} + \frac{1}{\beta_{\text{is}}^2} \|g\|_{L^2(\Omega)}. \tag{4.14}$$

Applying (4.12) gives

$$\|\boldsymbol{f}\|_{H^{-1}(\Omega)} = \sup_{\boldsymbol{v} \in V \setminus \{\boldsymbol{0}\}} \frac{\langle \boldsymbol{f}, \boldsymbol{v} \rangle_{V',V}}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)}} = \sup_{\boldsymbol{v} \in V \setminus \{\boldsymbol{0}\}} \frac{(\nabla \boldsymbol{w}, \nabla \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, r)}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)}}$$

and using in addition that $(\nabla \cdot \boldsymbol{w}, c) = 0$ for all $c \in \mathbb{R}$ yields

$$\|g\|_{L^2(\Omega)} = \sup_{q \in L^2(\Omega) \setminus \{0\}} \frac{(g, q)}{\|q\|_{L^2(\Omega)}} = \sup_{q \in L^2(\Omega) \setminus \{0\}} \frac{(\nabla \cdot \boldsymbol{w}, q)}{\|q\|_{L^2(\Omega)}} = \sup_{q \in Q \setminus \{0\}} \frac{(\nabla \cdot \boldsymbol{w}, q)}{\|q\|_{L^2(\Omega)}}.$$

Adding (4.13) and (4.14) leads to

$$\begin{aligned}
&\|\nabla \boldsymbol{w}\|_{L^2(\Omega)} + \|r\|_{L^2(\Omega)} \\
&\le \left(1 + \frac{2}{\beta_{\text{is}}}\right) \sup_{\boldsymbol{v} \in V \setminus \{\boldsymbol{0}\}} \frac{(\nabla \boldsymbol{w}, \nabla \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, r)}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)}} + \left(\frac{1}{\beta_{\text{is}}} + \frac{1}{\beta_{\text{is}}^2}\right) \sup_{q \in Q \setminus \{0\}} \frac{(\nabla \cdot \boldsymbol{w}, q)}{\|q\|_{L^2(\Omega)}} \\
&\le \left(\left(1 + \frac{2}{\beta_{\text{is}}}\right) + \frac{1}{\beta_{\text{is}}}\left(1 + \frac{1}{\beta_{\text{is}}}\right)\right) \sup_{\substack{(\boldsymbol{v},q) \in (V \times Q) \\ (\boldsymbol{v},q) \ne (\boldsymbol{0},0)}} \frac{(\nabla \boldsymbol{w}, \nabla \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, r) + (\nabla \cdot \boldsymbol{w}, q)}{\left(\|\nabla \boldsymbol{v}\|_{L^2(\Omega)} + \|q\|_{L^2(\Omega)}\right)},
\end{aligned}$$

where in the last step it was used that

$$\sup_{\boldsymbol{v} \in V \setminus \{\boldsymbol{0}\}} \frac{(\nabla \boldsymbol{w}, \nabla \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, r)}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)}} \le \sup_{\substack{(\boldsymbol{v},q) \in (V \times Q) \\ (\boldsymbol{v},q) \ne \{\boldsymbol{0},0\}}} \frac{(\nabla \boldsymbol{w}, \nabla \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, r) + (\nabla \cdot \boldsymbol{w}, q)}{\left(\|\nabla \boldsymbol{v}\|_{L^2(\Omega)} + \|q\|_{L^2(\Omega)}\right)}$$

because for $q = 0$ there already holds the equal sign. The same argument is applied to the other supremum. Since $(\boldsymbol{w}, r) \in V \times Q$ was arbitrary, the inf-sup condition (4.11) follows from this estimate. ■

Remark 4.12 (On the Balance of Different Parts of the Source Term in the Momentum Equation) Let $\boldsymbol{f} \in L^2(\Omega)$. According to Theorem 3.168, the right-hand side admits a Helmholtz decomposition of the form $\boldsymbol{f} = P_{\text{helm}}\boldsymbol{f} + \nabla r$, with $P_{\text{helm}}\boldsymbol{f} \in H_{\text{div}}(\Omega)$ and $\nabla r \in (H_{\text{div}}(\Omega))^{\perp}$. Inserting this decomposition in the momentum balance of the Stokes equations (4.1) yields

$$-\Delta\boldsymbol{u} + \nabla p = P_{\text{helm}}\boldsymbol{f} + \nabla r. \tag{4.15}$$

Assume that $\Delta\boldsymbol{u}, \nabla p \in L^2(\Omega)$, then it is even $\Delta\boldsymbol{u} \in H_{\text{div}}(\Omega)$. This property can be checked by a direct calculation, utilizing that weak derivatives can be always interchanged, see Evans (2010, Sect. 5.2.3, Theorem 1), and using $\nabla \cdot \boldsymbol{u} = 0$

$$\begin{aligned} -\nabla \cdot \Delta\boldsymbol{u} &= \partial_x\left(\partial_{xx}u_1 + \partial_{yy}u_1\right) + \partial_y\left(\partial_{xx}u_2 + \partial_{yy}u_2\right) \\ &= \partial_{xx}\left(\partial_x u_1 + \partial_y u_2\right) + \partial_{yy}\left(\partial_x u_1 + \partial_y u_2\right) = 0, \end{aligned}$$

and analogously in three dimensions.

Consider now not only the momentum balance but the complete boundary value problem. Inserting the decomposition of $\boldsymbol{f}$ in the weak formulation (4.2) and assuming that the solution is sufficiently smooth gives, applying integration by parts,

$$-(\Delta\boldsymbol{u}, \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, p) = (P_{\text{helm}}\boldsymbol{f}, \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, r) \quad \forall\ \boldsymbol{v} \in V. \tag{4.16}$$

It can be seen that the irrotational forces ∇r are balanced completely by the pressure. Already the stability estimate (4.9) shows that irrotational forces do not possess an impact on the velocity. Altogether, one finds that if $\boldsymbol{f}$ is changed to $\boldsymbol{f} + \nabla r$, then the pressure solution of the Stokes equations changes to $p + r$. Likewise, divergence-free forces are balanced by the velocity, more precisely by $-\Delta\boldsymbol{u}$. Thus, there is a separate balance of irrotational and divergence-free forces.

Note that for the integration by parts leading to (4.16) the respective boundary integrals have to vanish. This situation is given if the test space is $V = H_0^1(\Omega)$, i.e., if the Stokes equations are equipped with Dirichlet conditions on the whole boundary. There is not such a clear separation of the impact of different kinds of the forces in the case of other boundary conditions, in particular, in the case of boundary conditions which involve also the pressure, like the do-nothing condition (2.37). □

4.2 Finite Element Error Analysis

Remark 4.13 (The Finite Element Formulation) Let V^h be a velocity finite element space and let Q^h be a pressure finite element space. The finite element discretization of the Stokes equations (4.2) reads as follows: Let $\boldsymbol{f}$ be given, find $(\boldsymbol{u}^h, p^h) \in V^h \times Q^h$ such that

$$\begin{aligned} a^h\left(\boldsymbol{u}^h, \boldsymbol{v}^h\right) + b^h\left(\boldsymbol{v}^h, p^h\right) &= \sum_{K\in\mathcal{T}^h} \int_K \left(\boldsymbol{f}\boldsymbol{v}^h\right)(\boldsymbol{x})\, d\boldsymbol{x} \ \forall\ \boldsymbol{v}^h \in V^h, \\ b^h\left(\boldsymbol{u}^h, q^h\right) &= 0 \qquad\qquad\qquad\qquad\quad \forall\ q^h \in Q^h, \end{aligned} \tag{4.17}$$

with

$$a^h\left(\boldsymbol{v}^h, \boldsymbol{w}^h\right) = \sum_{K\in\mathcal{T}^h} \left(\nabla \boldsymbol{v}^h, \nabla \boldsymbol{w}^h\right)_K, \quad b^h\left(\boldsymbol{v}^h, q^h\right) = -\sum_{K\in\mathcal{T}^h} \left(\nabla\cdot \boldsymbol{v}^h, q^h\right)_K, \tag{4.18}$$

and the right-hand side $\boldsymbol{f}$ is assumed to be sufficiently smooth such that the right-hand side of the momentum equation in (4.17) is well defined. □

Theorem 4.14 (Existence and Uniqueness of a Solution of the Finite Element Stokes Equations) *Let* $(a^h(\cdot,\cdot))^{1/2}$ *define a norm in* V^h *and let the discrete inf-sup condition* (3.51) *hold. Then,* (4.17) *has a unique solution.*

Proof If $(a^h(\cdot,\cdot))^{1/2}$ is a norm in V^h, then this bilinear form is V^h-elliptic and in particular V^h_{div}-elliptic. Then, the proof is performed analogously to the proof of Theorem 4.6. ■

Remark 4.15 (Goal and Approach) The goal of the finite element error analysis consists in getting information on the order of convergence of the finite element solution to the solution of the weak problem in norms of interest. To this end, families of triangulations $\{\mathcal{T}^h\}$ with corresponding finite element spaces $\{V^h \times Q^h\}$ are considered. A general approach of obtaining finite element error estimates consists in the following steps, e.g., see the proof of Theorem 4.21:

- derive an equation or an inequality for the considered norm of the error,
- modify the equation or the inequality in such a way that approximation errors to the finite element spaces appear,
- estimate the considered norm of the error by constants times best approximation errors.

In the second step, one has usually to add and subtract terms in a clever way. The best approximation errors are independent of the considered problem. They can be estimated by interpolation errors. Interpolation error estimates are known from the general theory of finite element methods, see Appendix C. □

4.2.1 Conforming Inf-Sup Stable Pairs of Finite Element Spaces

Remark 4.16 (Conforming Inf-Sup Stable Finite Element Spaces) This section studies conforming inf-sup stable finite element spaces, i.e., besides the discrete inf-sup condition (3.51) it holds $V^h \subset V$ and $Q^h \subset Q$. The bilinear forms are identical to the forms of the continuous problem, i.e., it is $a^h(\cdot,\cdot) = a(\cdot,\cdot)$ and $b^h(\cdot,\cdot) = b(\cdot,\cdot)$. □

Corollary 4.17 (Unique Solvability of the Finite Element Problem) *Let* V^h *and* Q^h *be conforming finite element spaces that satisfy the discrete inf-sup condition* (3.51)*. Then, the finite element problem* (4.17) *has a unique solution.*

Proof Based on Theorem 4.14, it is sufficient to show that $(a(\cdot,\cdot))^{1/2}$ defines a norm in V^h. This property is given, because $(a(\cdot,\cdot))^{1/2}$ defines a norm in V, see (4.6), it defines also a norm in every subspace $V^h \subset V$. ■

Lemma 4.18 (Stability of the Finite Element Solution) *Let $V^h \times Q^h$ be a pair of inf-sup stable finite element spaces. Then, the solution of* (4.17) *fulfills*

$$\left\|\nabla \boldsymbol{u}^h\right\|_{L^2(\Omega)} \le \|\boldsymbol{f}\|_{H^{-1}(\Omega)}, \quad \left\|p^h\right\|_{L^2(\Omega)} \le \frac{2}{\beta_{\text{is}}^h}\|\boldsymbol{f}\|_{H^{-1}(\Omega)}.$$

Proof The proof follows the lines of the proof of Lemma 4.7. ■

Remark 4.19 (Reduction to a Problem in the Space of Discretely Divergence-Free Functions) From the abstract theory of linear saddle point problems, Sect. 3.1, Remark 3.11, it follows that the finite element approximation of the velocity can be computed by solving the following problem: Find $\boldsymbol{u}^h \in V_{\text{div}}^h$ such that

$$a\left(\boldsymbol{u}^h, \boldsymbol{v}^h\right) = \left(\nabla \boldsymbol{u}^h, \nabla \boldsymbol{v}^h\right) = \langle \boldsymbol{f}, \boldsymbol{v}^h\rangle_{V',V} \quad \forall\ \boldsymbol{v}^h \in V_{\text{div}}^h. \tag{4.19}$$

□

4.2.1.1 The Case $V_{\text{div}}^h \not\subset V_{\text{div}}$

Remark 4.20 (The Case $V_{\text{div}}^h \not\subset V_{\text{div}}$) This case applies for most pairs of inf-sup stable finite element spaces. □

Theorem 4.21 (Finite Element Error Estimate for the $L^2(\Omega)$ Norm of the Gradient of the Velocity) *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, be a bounded domain with polyhedral and Lipschitz continuous boundary and let $(\boldsymbol{u}, p) \in V \times Q$ be the unique solution of the Stokes problem* (4.2)*. Assume that this problem is discretized with inf-sup stable conforming finite element spaces $V^h \times Q^h$ and denote by $\boldsymbol{u}^h \in V_{\text{div}}^h$ the velocity solution. Then, the following error estimate holds*

$$\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)} \le 2 \inf_{\boldsymbol{v}^h \in V_{\text{div}}^h} \left\|\nabla(\boldsymbol{u}-\boldsymbol{v}^h)\right\|_{L^2(\Omega)} + \inf_{q^h \in Q^h} \left\|p-q^h\right\|_{L^2(\Omega)}. \tag{4.20}$$

Proof The proof starts by formulating the error equation. Since $V_{\text{div}}^h \subset V$, functions from V_{div}^h can be used as test function in the continuous Stokes equations (4.3), which is equivalent to (4.2). Using that the velocity solution of the continuous equation is weakly divergence-free, one obtains by subtracting (4.19) from (4.3)

$$\left(\nabla(\boldsymbol{u}-\boldsymbol{u}^h), \nabla \boldsymbol{v}^h\right) - \left(\nabla \cdot \boldsymbol{v}^h, p\right) = 0 \quad \forall\ \boldsymbol{v}^h \in V_{\text{div}}^h. \tag{4.21}$$

The pressure term appears since in general $V_{\text{div}}^h \not\subset V_{\text{div}}$. Equation (4.21) is the error equation.

Now, the pressure term will be modified such that an approximation term with respect to the pressure is obtained. Observing that $\left(\nabla\cdot\boldsymbol{v}^h, q^h\right) = 0$ for all $q^h \in Q^h$ leads to

$$\left(\nabla(\boldsymbol{u}-\boldsymbol{u}^h), \nabla\boldsymbol{v}^h\right) - \left(\nabla\cdot\boldsymbol{v}^h, p-q^h\right) = 0 \quad \forall\ \boldsymbol{v}^h \in V_{\mathrm{div}}^h,\ \forall\ q^h \in Q^h. \tag{4.22}$$

In the next step, the approximation error for the velocity will be introduced. To this end, the error is decomposed into

$$\boldsymbol{u}-\boldsymbol{u}^h = \left(\boldsymbol{u}-I^h\boldsymbol{u}\right) - \left(\boldsymbol{u}^h - I^h\boldsymbol{u}\right) = \boldsymbol{\eta} - \boldsymbol{\phi}^h.$$

Here, $I^h\boldsymbol{u}$ denotes an arbitrary interpolant of $\boldsymbol{u}$ in V_{div}^h. Hence, $\boldsymbol{\eta}$ is just an approximation error which depends only on the finite element space V_{div}^h but not on the considered equation. The goal is to estimate $\boldsymbol{\phi}^h \in V_{\mathrm{div}}^h$ by approximation errors as well. To this end, this decomposition is inserted in (4.22) and the test function $\boldsymbol{v}^h = \boldsymbol{\phi}^h$ is chosen. It follows that

$$\left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 = \left(\nabla\boldsymbol{\phi}^h, \nabla\boldsymbol{\phi}^h\right) = \left(\nabla\boldsymbol{\eta}, \nabla\boldsymbol{\phi}^h\right) - \left(\nabla\cdot\boldsymbol{\phi}^h, p-q^h\right) \quad \forall\ q^h \in Q^h. \tag{4.23}$$

Now, the terms on the right-hand side are estimated using the Cauchy–Schwarz inequality (A.10)

$$\left|\left(\nabla\boldsymbol{\eta}, \nabla\boldsymbol{\phi}^h\right)\right| \le \|\nabla\boldsymbol{\eta}\|_{L^2(\Omega)} \left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}$$

and, using in addition (3.170),

$$\begin{aligned} \left|-\left(\nabla\cdot\boldsymbol{\phi}^h, p-q^h\right)\right| &\le \left\|p-q^h\right\|_{L^2(\Omega)} \left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \\ &\le \left\|p-q^h\right\|_{L^2(\Omega)} \left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}. \end{aligned} \tag{4.24}$$

Inserting these estimates in (4.23) and dividing by $\left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \neq 0$ yields

$$\left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \le \|\nabla\boldsymbol{\eta}\|_{L^2(\Omega)} + \left\|p-q^h\right\|_{L^2(\Omega)}.$$

In the case that $\left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} = 0$, this estimate trivially holds.

With the triangle inequality it follows that

$$\begin{aligned} \left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)} &\le \left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} + \|\nabla\boldsymbol{\eta}\|_{L^2(\Omega)} \\ &\le 2\,\|\nabla\boldsymbol{\eta}\|_{L^2(\Omega)} + \left\|p-q^h\right\|_{L^2(\Omega)} \end{aligned}$$

for all $I^h\boldsymbol{u} \in V_{\mathrm{div}}^h$ and for all $q^h \in Q^h$, from what (4.20) follows. ■

Remark 4.22 (On Error Estimates (4.20)*)* The error in the V norm of the velocity is estimated with (4.20) by best approximation errors for both, the velocity and the pressure. The occurrence of the best approximation error for the velocity is natural, since the velocity finite element space has certainly an impact on the error. Inspecting the proof of Theorem 4.21, one finds that the reason for the appearance of the best approximation error for the pressure is the property $V_{\text{div}}^h \not\subset V_{\text{div}}$. Otherwise, the second term on the left-hand side of (4.3) would vanish and the pressure would have been eliminated from the estimate. The case $V_{\text{div}}^h \subset V_{\text{div}}$ will be studied in more detail in Sect. 4.2.1.2. □

Remark 4.23 (Conditions on the Domain) A polyhedral domain is assumed for the reason that no error is committed in approximating the boundary with the triangulations. The assumption of the Lipschitz continuity of the boundary is necessary, e.g., for the validity of Sobolev imbeddings on the considered domain. A domain with Lipschitz boundary lies only on one side of the boundary, see Adams (1975, p. 67). Under this condition, a polyhedral (polygonal) domain in two dimensions is a Lipschitz domain. However, the same implication does not hold for a polyhedral domain in three dimensions. For instance, if the domain is build of two bricks that are laying on each other like in Fig. 4.1, then the boundary is not Lipschitz continuous where the edge of one brick meets the edge of the other brick, see Král and Wendland (1986), Verchota and Vogel (2003) for details. □

Corollary 4.24 (Finite Element Error Estimate for the $L^2(\Omega)$ Norm of the Divergence of the Velocity) *Let the assumptions of Theorem 4.21 be fulfilled, then*

$$\left\|\nabla \cdot \boldsymbol{u}^h\right\|_{L^2(\Omega)} \le 2 \inf_{\boldsymbol{v}^h \in V_{\text{div}}^h} \left\|\nabla(\boldsymbol{u} - \boldsymbol{v}^h)\right\|_{L^2(\Omega)} + \inf_{q^h \in Q^h} \left\|p - q^h\right\|_{L^2(\Omega)}. \tag{4.25}$$

Proof The estimate follows easily from $\nabla \cdot \boldsymbol{u} = 0$, (3.170), and the estimate (4.20)

$$\left\|\nabla \cdot \boldsymbol{u}^h\right\|_{L^2(\Omega)} = \left\|\nabla \cdot (\boldsymbol{u} - \boldsymbol{u}^h)\right\|_{L^2(\Omega)} \le \left\|\nabla(\boldsymbol{u} - \boldsymbol{u}^h)\right\|_{L^2(\Omega)}.$$

■

Theorem 4.25 (Finite Element Error Estimate for the $L^2(\Omega)$ Norm of the Pressure) *Let the assumption of Theorem 4.21 be satisfied. Then the following*

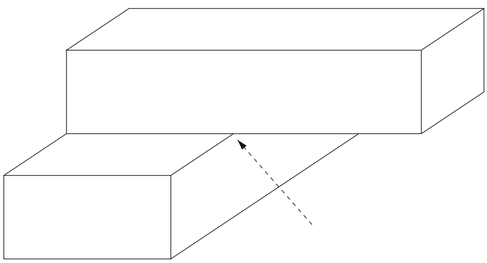

Fig. 4.1 Polyhedral domain in three dimensions that is not Lipschitz continuous (at the corner where the arrow points to)

error estimate holds

$$\left\|p - p^h\right\|_{L^2(\Omega)} \leq \frac{2}{\beta_{\mathrm{is}}^h} \inf_{\boldsymbol{v}^h \in V_{\mathrm{div}}^h} \left\|\nabla(\boldsymbol{u} - \boldsymbol{v}^h)\right\|_{L^2(\Omega)} + \left(1 + \frac{2}{\beta_{\mathrm{is}}^h}\right) \inf_{q^h \in Q^h} \left\|p - q^h\right\|_{L^2(\Omega)}. \tag{4.26}$$

Proof Let $q^h \in Q^h$ be arbitrary, then one gets with the triangle inequality

$$\left\|p - p^h\right\|_{L^2(\Omega)} \leq \left\|p - q^h\right\|_{L^2(\Omega)} + \left\|p^h - q^h\right\|_{L^2(\Omega)}.$$

Replacing the right-hand side of momentum equation of the finite element Stokes problem (4.17) by the left-hand side of the the momentum equation of the continuous Stokes problem (4.2) for $\boldsymbol{v}^h \in V^h$ gives

$$\begin{aligned} b\left(\boldsymbol{v}^h, p^h - q^h\right) &= -a\left(\boldsymbol{u}^h, \boldsymbol{v}^h\right) + \langle \boldsymbol{f}, \boldsymbol{v}^h\rangle_{V',V} - b\left(\boldsymbol{v}^h, q^h\right) \\ &= a\left(\boldsymbol{u} - \boldsymbol{u}^h, \boldsymbol{v}^h\right) + b\left(\boldsymbol{v}^h, p - q^h\right) \quad \forall\ \boldsymbol{v}^h \in V^h, \forall\ q^h \in Q^h. \end{aligned}$$

With the inf-sup condition (3.51), the Cauchy–Schwarz inequality (A.10), and (3.170), it follows now that

$$\begin{aligned} &\left\|p^h - q^h\right\|_{L^2(\Omega)} \\ &\leq \frac{1}{\beta_{\mathrm{is}}^h} \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{b\left(\boldsymbol{v}^h, p^h - q^h\right)}{\|\nabla \boldsymbol{v}^h\|_{L^2(\Omega)}} \\ &= \frac{1}{\beta_{\mathrm{is}}^h} \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{a\left(\boldsymbol{u} - \boldsymbol{u}^h, \boldsymbol{v}^h\right) + b\left(\boldsymbol{v}^h, p - q^h\right)}{\|\nabla \boldsymbol{v}^h\|_{L^2(\Omega)}} \\ &\leq \frac{1}{\beta_{\mathrm{is}}^h} \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{\left\|\nabla\left(\boldsymbol{u} - \boldsymbol{u}^h\right)\right\|_{L^2(\Omega)} \left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)} + \left\|p - q^h\right\|_{L^2(\Omega)} \left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)}}{\|\nabla \boldsymbol{v}^h\|_{L^2(\Omega)}} \\ &= \frac{1}{\beta_{\mathrm{is}}^h} \left(\left\|\nabla\left(\boldsymbol{u} - \boldsymbol{u}^h\right)\right\|_{L^2(\Omega)} + \left\|p - q^h\right\|_{L^2(\Omega)}\right) \quad \forall\ q^h \in Q^h. \end{aligned}$$

Inserting the error estimate (4.20) for the velocity yields the error estimate for the pressure. ■

Remark 4.26 (Error of the Velocity in the $L^2(\Omega)$ Norm) A simple error estimate of the velocity error in the $L^2(\Omega)$ can be obtained with the Poincaré inequality (A.12)

$$\left\|\boldsymbol{u} - \boldsymbol{u}^h\right\|_{L^2(\Omega)} \leq C \left\|\nabla(\boldsymbol{u} - \boldsymbol{u}^h)\right\|_{L^2(\Omega)}$$

and the application of (4.20). However, the resulting estimate is not optimal. The derivation of an optimal error estimate for $\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{L^2(\Omega)}$ requires additional assumptions and analytical tools. □

Remark 4.27 (Regular Dual Stokes Problem) To obtain an estimate for the velocity error in $L^2(\Omega)$, the classical argument by Aubin (1967) and Nitsche (1968) is applied. To this end, a dual problem of the Stokes equations has to be considered. The dual or adjoint operator is obtained by bringing the test function of the weak formulation of the Stokes equations (4.3) in the place of the solution and vice versa. Using the symmetry of the viscous term and changing position of terms, one gets from the left-hand side of (4.3)

$$(\nabla \boldsymbol{v}, \nabla \boldsymbol{u}) + (\nabla \cdot \boldsymbol{u}, q) - (\nabla \cdot \boldsymbol{v}, p).$$

Now, the test function is replaced by the solution of the dual problem $\left(\boldsymbol{\phi}_{\hat{f}}, \xi_{\hat{f}}\right)$ and the solution $(\boldsymbol{u}, p)$ by the test function, leading to the following left-hand side of the weak form of the dual problem

$$\left(\nabla \boldsymbol{\phi}_{\hat{f}}, \nabla \boldsymbol{v}\right) + \left(\nabla \cdot \boldsymbol{v}, \xi_{\hat{f}}\right) - \left(\nabla \cdot \boldsymbol{\phi}_{\hat{f}}, q\right).$$

Using integration by parts one gets the strong form of the dual Stokes problem for a velocity in V_{div}: For given $\hat{\boldsymbol{f}} \in L^2(\Omega)$, find $(\boldsymbol{\phi}_{\hat{f}}, \xi_{\hat{f}}) \in V \times Q$ such that

$$\begin{aligned} -\Delta \boldsymbol{\phi}_{\hat{f}} - \nabla \xi_{\hat{f}} &= \hat{\boldsymbol{f}} \text{ in } \Omega, \\ \nabla \cdot \boldsymbol{\phi}_{\hat{f}} &= 0 \text{ in } \Omega. \end{aligned} \tag{4.27}$$

Note that in the general dual problem, the right-hand side of the divergence constraint might be different than zero.

Problem (4.27) is said to be regular, if the mapping

$$\left(\boldsymbol{\phi}_{\hat{f}}, \xi_{\hat{f}}\right) \mapsto -\Delta \boldsymbol{\phi}_{\hat{f}} - \nabla \xi_{\hat{f}} \tag{4.28}$$

is an isomorphism from $\left(H^2(\Omega) \cap V\right) \times \left(H^1(\Omega) \cap Q\right)$ onto $L^2(\Omega)$. That means, in comparison with the Stokes equations, the higher regularity conditions $\boldsymbol{\phi}_{\hat{f}} \in H^2(\Omega)$ and $\xi_{\hat{f}} \in H^1(\Omega)$ are required. It can be proved that this regularity is given, e.g., for bounded convex polyhedral domains in two and three dimensions, see Kellogg and Osborn (1976), Dauge (1989). □

Theorem 4.28 (Finite Element Error Estimate for the $L^2(\Omega)$ Norm of the Velocity) *Let the assumption of Theorem 4.21 hold and let* $(\boldsymbol{\phi}_{\hat{f}}, \xi_{\hat{f}})$ *be the solution* *of* (4.27)*. Then, the* $L^2(\Omega)$ *error of the velocity can be estimated as follows*

$$\begin{aligned}\|\boldsymbol{u}-\boldsymbol{u}^h\|_{L^2(\Omega)} &\le \left(\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\|_{L^2(\Omega)} + \inf_{q^h\in Q^h}\|p-q^h\|_{L^2(\Omega)}\right)\\ &\quad\times \sup_{\hat{\boldsymbol{f}}\in L^2(\Omega)\setminus\{\boldsymbol{0}\}} \frac{1}{\left\|\hat{\boldsymbol{f}}\right\|_{L^2(\Omega)}}\left[\inf_{\boldsymbol{\phi}^h\in V^h_{\mathrm{div}}}\left\|\nabla\left(\boldsymbol{\phi}_{\hat{f}}-\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)}\right.\\ &\quad\left. + \inf_{r^h\in Q^h}\left\|\xi_{\hat{f}}-r^h\right\|_{L^2(\Omega)}\right]. \qquad (4.29)\end{aligned}$$

Proof It is, by definition,

$$\|\boldsymbol{u}-\boldsymbol{u}^h\|_{L^2(\Omega)} = \sup_{\hat{\boldsymbol{f}}\in L^2(\Omega)\setminus\{\boldsymbol{0}\}} \frac{\left(\hat{\boldsymbol{f}},\boldsymbol{u}-\boldsymbol{u}^h\right)}{\left\|\hat{\boldsymbol{f}}\right\|_{L^2(\Omega)}}. \qquad (4.30)$$

The weak form of the dual problem (4.27) is: Find $(\boldsymbol{\phi}_{\hat{f}},\xi_{\hat{f}})\in V\times Q$ such that

$$\begin{aligned}\left(\nabla\boldsymbol{v},\nabla\boldsymbol{\phi}_{\hat{f}}\right)+\left(\nabla\cdot\boldsymbol{v},\xi_{\hat{f}}\right) &= \left(\hat{\boldsymbol{f}},\boldsymbol{v}\right) &&\forall\ \boldsymbol{v}\in V,\\ \left(\nabla\cdot\boldsymbol{\phi}_{\hat{f}},q\right) &= 0 &&\forall\ q\in Q.\end{aligned}$$

Choosing $\boldsymbol{v}=\boldsymbol{u}-\boldsymbol{u}^h\in V$ gives for the numerator of (4.30)

$$\left(\hat{\boldsymbol{f}},\boldsymbol{u}-\boldsymbol{u}^h\right) = \left(\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right),\nabla\boldsymbol{\phi}_{\hat{f}}\right)+\left(\nabla\cdot\left(\boldsymbol{u}-\boldsymbol{u}^h\right),\xi_{\hat{f}}\right). \qquad (4.31)$$

The aim consists now in adding terms to this equation such that approximation errors are obtained. For all $\boldsymbol{\phi}^h\in V^h_{\mathrm{div}}\subset V$ and $q^h\in Q^h$ it holds, using the weak form of the Stokes problem (4.2) and the finite element problem (4.19),

$$\begin{aligned}\left(\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right),\nabla\boldsymbol{\phi}^h\right) &= \left(\nabla\cdot\boldsymbol{\phi}^h,p\right)+\langle\boldsymbol{f},\boldsymbol{\phi}^h\rangle_{V',V}-\left(\nabla\cdot\boldsymbol{\phi}^h,p^h\right)-\langle\boldsymbol{f},\boldsymbol{\phi}^h\rangle_{V',V}\\ &= \left(\nabla\cdot\boldsymbol{\phi}^h,p\right) = \left(\nabla\cdot\boldsymbol{\phi}^h,p-q^h\right)\quad\forall\ q^h\in Q^h. \qquad (4.32)\end{aligned}$$

In addition, it is

$$\left(\nabla\cdot\boldsymbol{\phi}_{\hat{f}},p-q^h\right)=0\quad\forall\ q^h\in Q^h,$$

and

$$\left(\nabla\cdot\left(\boldsymbol{u}-\boldsymbol{u}^h\right),r^h\right)=0\quad\forall\ r^h\in Q^h,$$

since $Q^h \subset Q$. Inserting these terms in (4.31) leads to

$$\left(\hat{\boldsymbol{f}}, \boldsymbol{u}-\boldsymbol{u}^h\right) = \left(\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right), \nabla\left(\boldsymbol{\phi}_{\hat{f}}-\boldsymbol{\phi}^h\right)\right) + \left(\nabla\cdot\left(\boldsymbol{\phi}_{\hat{f}}-\boldsymbol{\phi}^h\right), p-q^h\right) + \left(\nabla\cdot\left(\boldsymbol{u}-\boldsymbol{u}^h\right), \xi_{\hat{f}}-r^h\right) \tag{4.33}$$

for all $\boldsymbol{\phi}^h \in V_{\mathrm{div}}^h$ and all $q^h, r^h \in Q^h$. The application of the Cauchy–Schwarz inequality (A.10) and of (3.170) yields

$$\begin{aligned}\left|\left(\hat{\boldsymbol{f}}, \boldsymbol{u}-\boldsymbol{u}^h\right)\right| &\le \left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)} \left\|\nabla\left(\boldsymbol{\phi}_{\hat{f}}-\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)} \\ &\quad + \left\|p-q^h\right\|_{L^2(\Omega)} \left\|\nabla\left(\boldsymbol{\phi}_{\hat{f}}-\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)} \\ &\quad + \left\|\xi_{\hat{f}}-r^h\right\|_{L^2(\Omega)} \left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)} \\ &\le \left(\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)} + \left\|p-q^h\right\|_{L^2(\Omega)}\right) \\ &\quad \times \left(\left\|\nabla\left(\boldsymbol{\phi}_{\hat{f}}-\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)} + \left\|\xi_{\hat{f}}-r^h\right\|_{L^2(\Omega)}\right)\end{aligned} \tag{4.34}$$

for all $\boldsymbol{\phi}^h \in V_{\mathrm{div}}^h$ and all $q^h, r^h \in Q^h$. Taking the infimum of all approximation errors gives (4.29). ■

Remark 4.29 (On the Dependency of the Error Bounds on the Discrete Inf-Sup Constant) One can estimate the best approximation error with respect to V_{div}^h in (4.20), (4.25), (4.26), and (4.29) with (3.65) giving a term with $\left(\beta_{\mathrm{is}}^h\right)^{-1}$. The obtained estimates are worst case estimates.

For finite element spaces where the local interpolation operator that preserves the discrete divergence, studied in Girault and Scott (2003), can be constructed, (3.71) can be used instead of (3.65). Then, the constants in the velocity estimates depend on the inverse of local discrete inf-sup constants, compare Remark 3.62. In contrast to the error estimates with respect to the velocity, the error bound (4.26) for the pressure depends always on $\left(\beta_{\mathrm{is}}^h\right)^{-1}$.

In all cases, if the local inf-sup constants or β_{is}^h depend on the mesh width, then an optimal order of convergence cannot be expected. □

Corollary 4.30 (Finite Element Error Estimates for Conforming Inf-Sup Stable Pairs of Finite Element Spaces) *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be a bounded domain with polyhedral and Lipschitz continuous boundary which is decomposed by a regular and quasi-uniform family of triangulations $\{\mathcal{T}^h\}$. Let $(\boldsymbol{u}, p)$ be the solution of the Stokes equations (4.2) with $\boldsymbol{u} \in H^{k+1}(\Omega) \cap V$ and $p \in H^k(\Omega) \cap Q$. Then for* *the inf-sup stable pairs of finite element spaces*

- P_k^{bubble}/P_k, $k = 1$ *(MINI element),*
- P_k/P_{k-1}, Q_k/Q_{k-1}, $k \geq 2$ *(Taylor–Hood element),*
- $P_2^{\text{bubble}}/P_1^{\text{disc}}$, $P_3^{\text{bubble}}/P_2^{\text{disc}}$, $P_2^{\text{BR}}/P_1^{\text{disc}}$, $Q_k/P_{k-1}^{\text{disc}}$, $k \geq 2$,

the following error estimates hold

$$\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)} \leq Ch^k \left(\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \|p\|_{H^k(\Omega)}\right), \tag{4.35}$$

$$\left\|\nabla\cdot\boldsymbol{u}^h\right\|_{L^2(\Omega)} \leq Ch^k \left(\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \|p\|_{H^k(\Omega)}\right), \tag{4.36}$$

$$\left\|p-p^h\right\|_{L^2(\Omega)} \leq Ch^k \left(\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \|p\|_{H^k(\Omega)}\right). \tag{4.37}$$

If the dual Stokes problem (4.27) *possesses a regular solution* $(\boldsymbol{\phi}_{\hat{f}}, \xi_{\hat{f}})$ *in the sense of Remark 4.27, then it holds in addition*

$$\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{L^2(\Omega)} \leq Ch^{k+1} \left(\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \|p\|_{H^k(\Omega)}\right). \tag{4.38}$$

The constants C depend either on the inverse of the discrete inf-sup constant β_{is}^h *or on the inverse of local inf-sup constants, compare Remark 4.29.*

Proof The best approximation errors in estimates (4.20), (4.25), (4.26), and (4.29) can be estimated by interpolation errors, since an interpolation error cannot be lower than the best approximation error. Then, estimates (4.35)–(4.37) follow directly from interpolation error estimates for finite element spaces, see (3.65) or (3.71) for the velocity.

The additional power of h in estimate (4.38) comes from the application of the interpolation estimates (3.65) or (3.71) and for the pressure to the second factor of (4.29) and the regularity of $(\boldsymbol{\phi}_{\hat{f}}, \xi_{\hat{f}})$

$$\begin{aligned}\inf_{\boldsymbol{\phi}^h\in V_{\text{div}}^h} \left\|\nabla\left(\boldsymbol{\phi}_{\hat{f}}-\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)} + \inf_{r^h\in Q^h}\left\|\xi_{\hat{f}}-r^h\right\|_{L^2(\Omega)} \\ \leq Ch\left(\left\|\boldsymbol{\phi}_{\hat{f}}\right\|_{H^2(\Omega)} + \left\|\xi_{\hat{f}}\right\|_{H^1(\Omega)}\right).\end{aligned} \tag{4.39}$$

Since (4.28) is an isomorphism from $\left(H^2(\Omega)\cap V\right)\times\left(H^1(\Omega)\cap Q\right)$ onto $L^2(\Omega)$ there is a constant C such that

$$\left\|\boldsymbol{\phi}_{\hat{f}}\right\|_{H^2(\Omega)} + \left\|\xi_{\hat{f}}\right\|_{H^1(\Omega)} \leq C\left\|\hat{\boldsymbol{f}}\right\|_{L^2(\Omega)}. \tag{4.40}$$

Inserting (4.39) and (4.40) in (4.29) proves (4.38). ∎

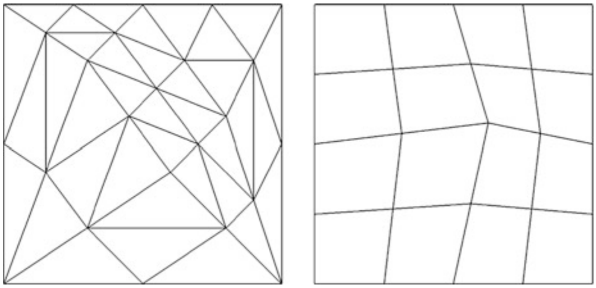

Fig. 4.2 Example 4.31. Initial grids (level 0)

Example 4.31 (Analytical Example Which Supports the Error Estimates (4.35)–(4.38)*)* Error estimates of form (4.35)–(4.38) are usually supported by using problems with a known (prescribed) solution and by measuring the errors to this solution on a sequence of subsequently refined grids. Here, Example D.3 will be considered, which is defined on the unit square. Initial grids (level 0) of the form presented in Fig. 4.2 were used in the simulations. On such unstructured or distorted grids, it is not likely to obtain superconvergence effects in the numerical results. The right-hand side was integrated with a quadrature rule of high order to reduce the influence of quadrature errors.

The errors in different norms are presented in Figs. 4.3, 4.4, 4.5, and 4.6.[1] It can be observed that in the most cases the orders of convergence predicted by the numerical analysis coincide with the orders of convergence in the numerical simulations. Generally, different discretizations of the same order show a similar accuracy. Only the pressure error obtained with $P_2^{\text{bubble}}/P_1^{\text{disc}}$, see Fig. 4.5, is much higher than with all other discretizations with second order velocity and first order pressure.

For the P_1^{bubble}/P_1 pair of finite element spaces, the order of convergence for the $L^2(\Omega)$ error of the pressure is higher by 0.5 than predicted by the analysis. □

Remark 4.32 (Local Mass Conservation for Discontinuous Pressure Approximations) An argument for using discontinuous pressure approximations, which can be found sometimes in the literature, is that one has at least a local (mesh cell by mesh cell) divergence-free finite element solution or local conservation of mass.

Consider meshes with affinely mapped grid cells. Since the piecewise constant functions are usually a subspace of a discontinuous pressure finite element space, one obtains from the second equation in (4.17)

$$0 = \sum_{K\in\mathcal{T}^h} \int_K \left(\nabla\cdot\boldsymbol{u}^h\right) q^h(\boldsymbol{x})\, d\boldsymbol{x} = \sum_{K\in\mathcal{T}^h} q^h \int_K \left(\nabla\cdot\boldsymbol{u}^h\right)(\boldsymbol{x})\, d\boldsymbol{x} \tag{4.41}$$

[1]2D graphics were plotted with Matplotlib, Hunter (2007), http://matplotlib.org.

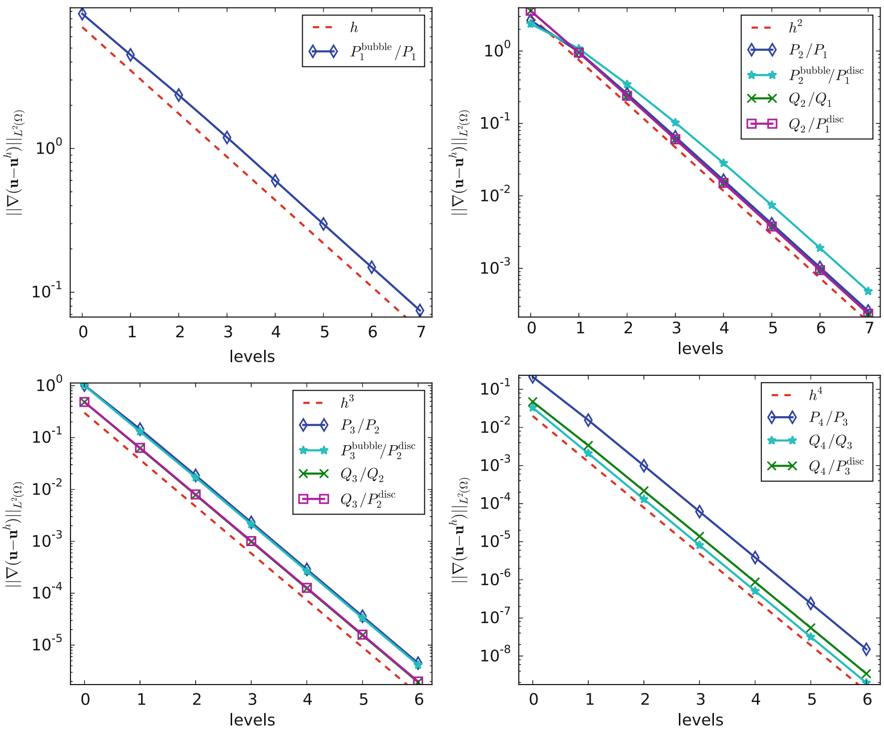

Fig. 4.3 Example 4.31. Convergence of the errors $\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)}$ for different discretizations with different orders k. In the *top right* and *bottom left* pictures, the *magenta line* is on *top* of the *green line*

for all $q^h \in P_0$ or $q^h \in Q_0$. Considering an arbitrary mesh cell K_1 and another arbitrary mesh cell $K_2 \neq K_1$. Then one can choose

$$q^h = \begin{cases} 1 & \text{in } K_1, \\ -\dfrac{|K_1|}{|K_2|} & \text{in } K_2, \\ 0 & \text{else.} \end{cases}$$

With this choice it is $q^h \in L_0^2(\Omega)$. One gets with (4.41)

$$\int_{K_2} \nabla \cdot \boldsymbol{u}^h(\boldsymbol{x})\, d\boldsymbol{x} = \frac{|K_2|}{|K_1|} \int_{K_1} \nabla \cdot \boldsymbol{u}^h(\boldsymbol{x})\, d\boldsymbol{x} \quad \forall\, K_2 \in \mathcal{T}^h.$$

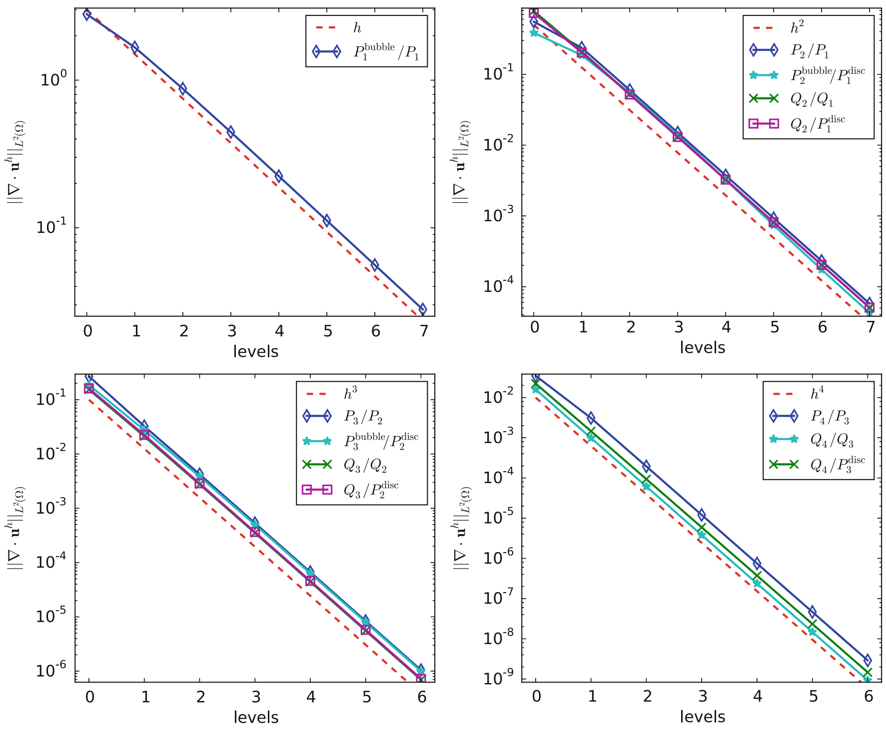

Fig. 4.4 Example 4.31. Convergence of the errors $\|\nabla \cdot \boldsymbol{u}^h\|_{L^2(\Omega)}$ for different discretizations with different orders k. In the *top right* and *bottom left* pictures, the *magenta line* is on *top* of the *green line*

It follows that

$$\begin{aligned}\int_\Omega \nabla \cdot \boldsymbol{u}^h(\boldsymbol{x})\, d\boldsymbol{x} &= \sum_{K\in\mathcal{T}^h} \int_K \nabla \cdot \boldsymbol{u}^h(\boldsymbol{x})\, d\boldsymbol{x} = \sum_{K\in\mathcal{T}^h} \frac{|K|}{|K_1|} \int_{K_1} \nabla \cdot \boldsymbol{u}^h(\boldsymbol{x})\, d\boldsymbol{x} \\ &= \frac{1}{|K_1|} \int_{K_1} \nabla \cdot \boldsymbol{u}^h(\boldsymbol{x})\, d\boldsymbol{x} \sum_{K\in\mathcal{T}^h} |K| \\ &= \frac{|\Omega|}{|K_1|} \int_{K_1} \nabla \cdot \boldsymbol{u}^h(\boldsymbol{x})\, d\boldsymbol{x}. \end{aligned} \tag{4.42}$$

Since by integration by parts one has

$$\int_\Omega \nabla \cdot \boldsymbol{u}^h(\boldsymbol{x})\, d\boldsymbol{x} = \int_{\partial\Omega} \boldsymbol{u}^h \cdot \boldsymbol{n}(\boldsymbol{s})\, ds = 0,$$

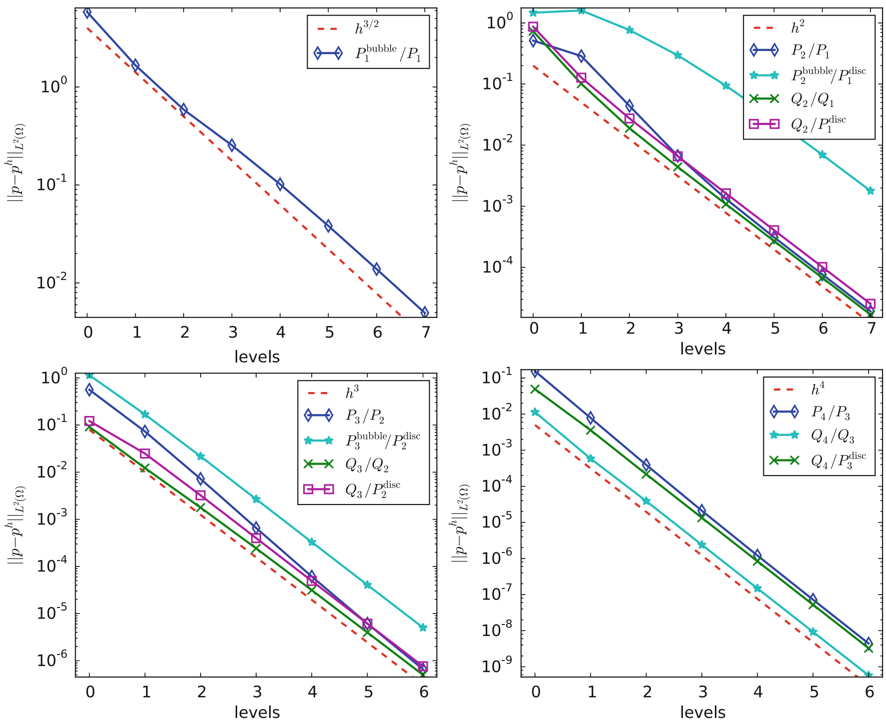

Fig. 4.5 Example 4.31. Convergence of the errors $\|p - p^h\|_{L^2(\Omega)}$ for different discretizations with different orders k

one concludes that the last factor on the right-hand side of (4.42) vanishes. Since K_1 was chosen to be arbitrary, one obtains the local mass conservation

$$\int_K \nabla \cdot \boldsymbol{u}^h(\boldsymbol{x})\, d\boldsymbol{x} = 0 \quad \forall\ K \in \mathcal{T}^h. \tag{4.43}$$

However, one can observe in Fig. 4.4 that despite the local mass conservation (4.43) for discretizations with discontinuous pressure approximations, the order of magnitude of the error $\|\nabla \cdot \boldsymbol{u}^h\|_{L^2(\Omega)}$ is the same for both, finite element discretizations with continuous and discontinuous pressure space. □

Remark 4.33 (Scaled Stokes Problem) A scaled Stokes problem of the form

$$\begin{aligned} -\nu \Delta \boldsymbol{u} + \nabla p &= \boldsymbol{f} \text{ in } \Omega, \\ \nabla \cdot \boldsymbol{u} &= 0 \text{ in } \Omega, \end{aligned} \tag{4.44}$$

$\nu > 0$, is sometimes of interest for academic purposes, since in (4.44), the viscous term is scaled in the same form as for the Navier–Stokes equations. Dividing the

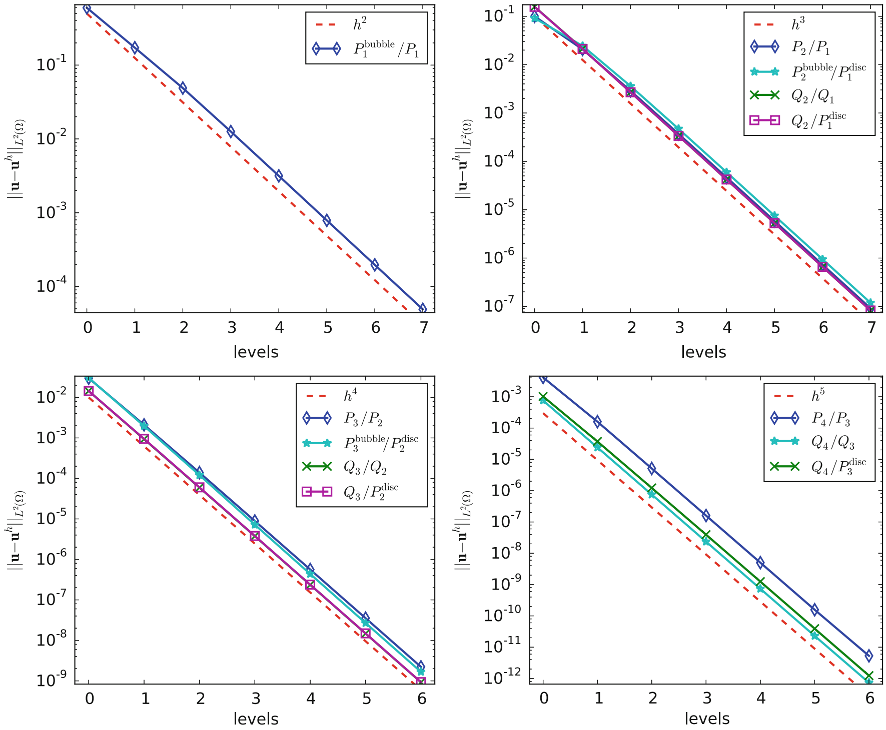

Fig. 4.6 Example 4.31. Convergence of the errors $\|\boldsymbol{u} - \boldsymbol{u}^h\|_{L^2(\Omega)}$ for different discretizations with different orders k

momentum equation in (4.44) by ν yields

$$\begin{aligned} -\Delta \boldsymbol{u} + \nabla \left(\frac{p}{\nu}\right) &= \frac{\boldsymbol{f}}{\nu} \text{ in } \Omega, \\ \nabla \cdot \boldsymbol{u} &= 0 \;\text{ in } \Omega, \end{aligned}$$

which is of the same form as the unscaled version (4.1) with a new pressure and a new source term. Now, the finite element error analysis can be applied in the same way as presented in this section, leading to the estimates of form (4.20), (4.25), (4.26), and (4.29), where the pressure terms (and in (4.29) the term with the dual pressure) are scaled with ν^{-1}. Consequently, one obtains also estimates of the form (4.35) and (4.36) with ν^{-1} in front of $\|p\|_{H^k(\Omega)}$

$$\begin{aligned} \left\|\nabla(\boldsymbol{u} - \boldsymbol{u}^h)\right\|_{L^2(\Omega)} &\le Ch^k \left(\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \nu^{-1}\|p\|_{H^k(\Omega)}\right), \\ \left\|\nabla \cdot \boldsymbol{u}^h\right\|_{L^2(\Omega)} &\le Ch^k \left(\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \nu^{-1}\|p\|_{H^k(\Omega)}\right). \end{aligned} \tag{4.45}$$

For the estimate of the $L^2(\Omega)$ error of the velocity, one considers again the dual Stokes problem (4.27). Since this problem is formulated with unit viscosity, neither $\boldsymbol{\phi}_{\hat{f}}$ nor $\xi_{\hat{f}}$ nor $\hat{\boldsymbol{f}}$ depend on ν. The error estimate is performed in the same way as in the proof of Theorem 4.28. Instead of (4.32), one obtains

$$\left(\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right),\nabla\boldsymbol{\phi}^h\right)=\frac{1}{\nu}\left(\nabla\cdot\boldsymbol{\phi}^h,p-q^h\right)\quad\forall\, q^h\in Q^h.$$

Then, the middle term on the right-hand side of (4.33) is scaled with ν^{-1} and one gets the scaling ν^{-1} in (4.34) in front of $\left\|p-q^h\right\|_{L^2(\Omega)}$. Finally, the error estimate

$$\begin{aligned}\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{L^2(\Omega)}\le{}& Ch^{k+1}\left(\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}+\nu^{-1}\|p\|_{H^k(\Omega)}\right)\\ &\times\left(\left\|\boldsymbol{\phi}_{\hat{f}}\right\|_{H^2(\Omega)}+\left\|\xi_{\hat{f}}\right\|_{H^1(\Omega)}\right)\end{aligned}\tag{4.46}$$

is derived.

Thus, for small values of ν, the term $\nu^{-1}\|p\|_{H^k(\Omega)}$ might dominate the right-hand side of all velocity error bounds.

In estimate (4.37) for the pressure, one has to scale also the term on the left-hand side with ν^{-1}. Rescaling this estimate leads to

$$\left\|p-p^h\right\|_{L^2(\Omega)}\le Ch^k\left(\nu\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}+\|p\|_{H^k(\Omega)}\right).\tag{4.47}$$

□

Example 4.34 (Scaled Stokes Problem) Again, the problem defined in Example D.3 is considered, see Example 4.31 for the simulations with the unscaled Stokes equations. From the estimates (4.45) and (4.46), one would expect that the velocity errors become large for small ν and then they scale linearly with ν^{-1}. In contrast, from (4.47) one has the expectation that the pressure error becomes large for large values of ν and then it scales linearly with ν.

Representative results for the second order Taylor–Hood pair of finite element spaces P_2/P_1 on the unstructured grid from Fig. 4.2 are presented in Fig. 4.7. The dependency of the velocity errors on ν^{-1} and the pressure error on ν is clearly visible. On coarse grids, also the linear dependencies on ν^{-1} and ν, respectively, can be observed. However, one can also see a higher order of decrease for the curves with large errors until they reach the curves for which a dependency on the value of ν cannot be observed. This decrease is higher by half an order for the velocity errors and by one order for the $L^2(\Omega)$ error of the pressure. To the best of our knowledge, there is no explanation for this behavior so far. □

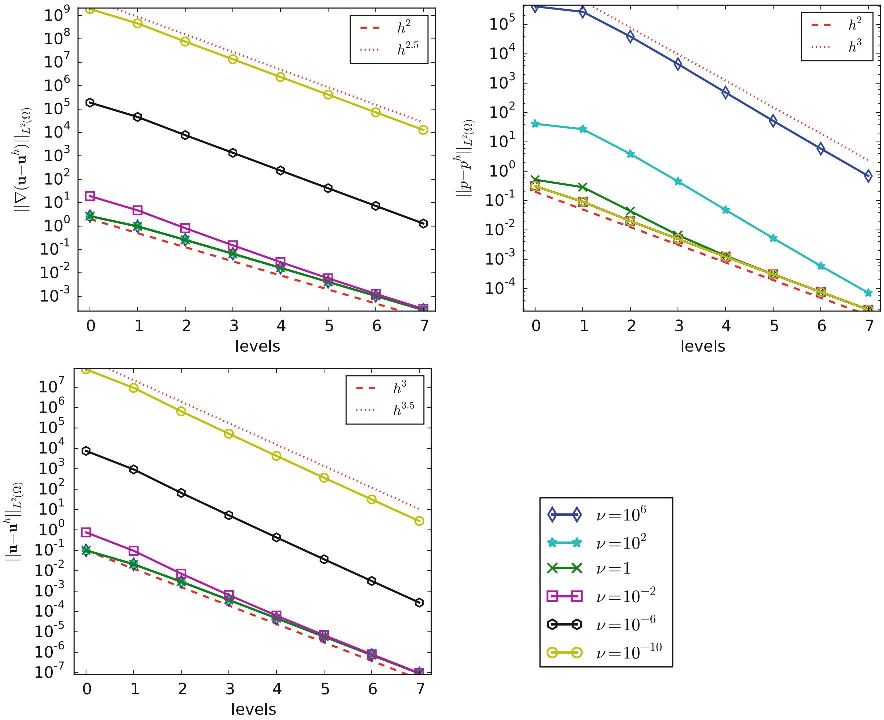

Fig. 4.7 Example 4.34. Convergence of the errors for the scaled Stokes problem and the P_2/P_1 pair of finite element spaces

4.2.1.2 The Case $V_{\mathrm{div}}^h \subset V_{\mathrm{div}}$

Remark 4.35 (Pairs of Finite Element Spaces with $V_{\mathrm{div}}^h \subset V_{\mathrm{div}}$*)* This section inspects the proofs of the error estimates from Sect. 4.2.1.1 under the condition that $V_{\mathrm{div}}^h \subset V_{\mathrm{div}}$. It turns out that some terms vanish. An important consequence is that the error estimates for the velocity do not depend any longer on the best approximation errors of the pressure finite element space. In addition, also a scaling of the viscous term as discussed in Remark 4.33 does not influence the velocity error estimates. The estimates in the case $V_{\mathrm{div}}^h \subset V_{\mathrm{div}}$ reflect the physics of the problem properly, in contrast to the estimates for the case $V_{\mathrm{div}}^h \not\subset V_{\mathrm{div}}$, compare Remark 4.12.

The most important pair of conforming inf-sup stable finite element spaces satisfying $V_{\mathrm{div}}^h \subset V_{\mathrm{div}}$ is the Scott–Vogelius pair of finite element spaces $P_k/P_{k-1}^{\mathrm{disc}}$, $k \geq d$, on special grids, see Sect. 3.6.3. □

Corollary 4.36 (Finite Element Error Estimates for the Velocity for Inf-Sup Stable Pairs of Finite Element Spaces with $V_{\mathrm{div}}^h \subset V_{\mathrm{div}}$) *Let the assumptions of Theorem 4.21 be fulfilled and consider an inf-sup stable pair of finite element* *spaces with* $V^h_{\mathrm{div}} \subset V_{\mathrm{div}}$*, then*

$$\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)} \le 2 \inf_{\boldsymbol{v}^h\in V^h_{\mathrm{div}}} \left\|\nabla(\boldsymbol{u}-\boldsymbol{v}^h)\right\|_{L^2(\Omega)} \tag{4.48}$$

and

$$\left\|\nabla\cdot\boldsymbol{u}^h\right\|_{L^2(\Omega)} = 0. \tag{4.49}$$

Proof The proof of (4.48) is performed in the same way as the proof of Theorem 4.21. Inspecting this proof for pairs of spaces with $V^h_{\mathrm{div}} \subset V_{\mathrm{div}}$, one finds that (4.24) equals zero since $\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} = 0$ for all $\boldsymbol{\phi}^h \in V^h_{\mathrm{div}}$.

Property (4.49) follows directly from the definition of V_{div}. ∎

Corollary 4.37 (Finite Element Error Estimate for the $L^2(\Omega)$ Norm of the Velocity for Inf-Sup Stable Pairs of Finite Element Spaces with $V^h_{\mathrm{div}} \subset V_{\mathrm{div}}$) *Let the assumptions of Theorem 4.28 be fulfilled. If for an inf-sup stable pair of finite element spaces* $V^h_{\mathrm{div}} \subset V_{\mathrm{div}}$*, then*

$$\begin{aligned}\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{L^2(\Omega)} &\le \left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}\\ &\times \sup_{\hat{\boldsymbol{f}}\in L^2(\Omega)} \frac{1}{\left\|\hat{\boldsymbol{f}}\right\|_{L^2(\Omega)}} \inf_{\boldsymbol{v}^h\in V^h_{\mathrm{div}}} \left\|\nabla\left(\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}-\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)}.\end{aligned} \tag{4.50}$$

Proof The proof proceeds in the same way as the proof of Theorem 4.28. In addition, one can use in (4.33) that $\nabla\cdot\left(\boldsymbol{\phi}^h-\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}\right) = 0$ and $\nabla\cdot\left(\boldsymbol{u}-\boldsymbol{u}^h\right) = 0$ in the weak sense. ∎

Remark 4.38 (Pairs of Finite Element Spaces with $V^h_{\mathrm{div}} \subset V_{\mathrm{div}}$*)* For pairs of finite element spaces with the property $V^h_{\mathrm{div}} \subset V_{\mathrm{div}}$, it follows from (4.48) and (4.49) that

$$\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)} \le Ch^k \left\|\boldsymbol{u}\right\|_{H^{k+1}(\Omega)}, \tag{4.51}$$

$$\left\|\nabla\cdot\boldsymbol{u}^h\right\|_{L^2(\Omega)} = 0, \tag{4.52}$$

$$\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{L^2(\Omega)} \le Ch^{k+1} \left\|\boldsymbol{u}\right\|_{H^{k+1}(\Omega)}. \tag{4.53}$$

These estimates are in particular true for the Scott–Vogelius spaces $P_k/P^{\mathrm{disc}}_{k-1}$, $k \ge d$, on barycentric-refined grids. □

Remark 4.39 (Scaled Stokes Problem) Considering a scaled Stokes problem of the form (4.44), one finds that the scaling ν^{-1} does not affect the velocity error estimates, in contrast to pairs of finite element spaces with $V^h_{\mathrm{div}} \not\subset V_{\mathrm{div}}$, see Remark 4.33. □

Example 4.40 (The Scott–Vogelius Pair of Finite Element Spaces P_2/P_1^{disc} on a Barycentric-Refined Grid for a No-flow Problem) In this example, the scaled Stokes problem (4.44) is considered in $\Omega = (0,1)^2$ and with the prescribed solution $\boldsymbol{u} = \mathbf{0}$ and the pressure

$$p(x,y) = 10\left((x-0.5)^3 y + (1-x)^2(y-0.5)^2 - \frac{1}{36}\right).$$

The Scott–Vogelius pair is known to satisfy the discrete inf-sup condition on barycentric-refined grids, see Remark 3.135. The grids were constructed as follows. The unit square was divided into two triangles by connecting the lower left and the upper right corner. This triangulation was uniformly refined once. Then a barycentric-refined grid as depicted in Fig. 3.9 was created, giving level 0. After having simulated the problem on this grid, the barycentric refinements were removed, the grid was uniformly refined once more, and again a barycentric refinement was applied, leading to level 1. This process was continued.

The results are presented in Fig. 4.8. The velocity error is always a small constant, independently of the value of ν. The increase of this constant is due to the increase

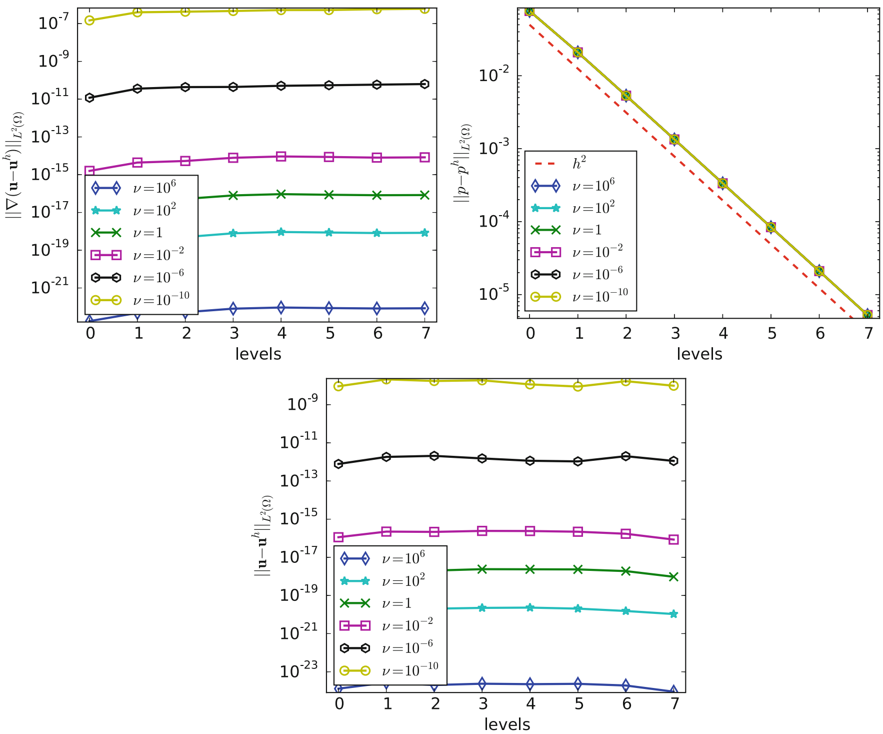

Fig. 4.8 Example 4.40. Convergence of the errors for the scaled Stokes problem and the P_2/P_1^{disc} pair of finite element spaces on a barycentric-refined grid

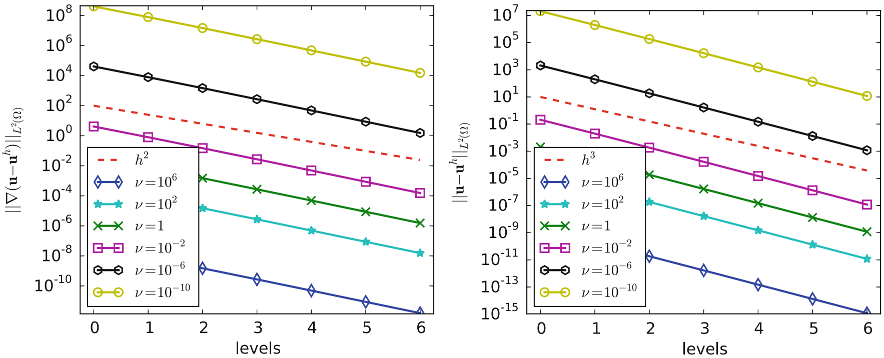

Fig. 4.9 Example 4.40. Convergence of the velocity errors for the scaled Stokes problem and the P_2/P_1 pair of finite element spaces

of the condition number of the linear saddle point problems for small ν. Hence, this increase reflects round-off errors of the solver. Since the first term on the right-hand side of estimate (4.47) vanishes, one expects that the pressure error in $L^2(\Omega)$ is independent of ν, which can be observed very well in the computational results.

As comparison, results obtained with the Taylor–Hood finite element P_2/P_1 computed on the irregular grid from Fig. 4.2 are depicted in Fig. 4.9. Even for moderate values of ν, the discrete velocity field is far away from being a no-flow field. The dependency of the velocity errors on the viscosity can be clearly observed in this simple example. This result reflects once more the potential impact of the pressure on the error of the velocity for pairs of spaces that do not satisfy $V_{\text{div}}^h \subset V_{\text{div}}$. Note that the order of convergence of both velocity errors is higher by 0.5 than predicted by the analysis. □

4.2.2 *The Stokes Projection*

Remark 4.41 (The Stokes Projection) Instead of using the best approximation error in V_{div}^h, some approaches of the finite element error analysis of the Navier–Stokes equations consider a concrete function in V_{div}^h, which is known to satisfy the required interpolation error estimates, e.g., see the proof of Theorem 7.35. A possible definition of a concrete function is the Stokes projection.

Given $(\boldsymbol{u}, p) \in V_{\text{div}} \times Q$, then the Stokes projection $(I_{\text{St}}^h \boldsymbol{u}, I_{\text{St}}^h p) \in V_{\text{div}}^h \times Q^h$ is defined to be the solution of

$$\begin{aligned} a\left(I_{\text{St}}^h \boldsymbol{u}, \boldsymbol{v}^h\right) + b\left(\boldsymbol{v}^h, I_{\text{St}}^h p\right) &= a\left(\boldsymbol{u}, \boldsymbol{v}^h\right) + b\left(\boldsymbol{v}^h, p\right) \ \forall\ \boldsymbol{v}^h \in V^h, \\ b\left(I_{\text{St}}^h \boldsymbol{u}, q^h\right) &= 0 \qquad\qquad\qquad\qquad\ \ \forall\ q^h \in Q^h, \end{aligned} \tag{4.54}$$

where $V^h \times Q^h \subset V \times Q$ are inf-sup stable finite element spaces. □

Lemma 4.42 (Stability of the Stokes Projection) *The Stokes projection depends stable on the solution of the problem*

$$\left\|\nabla I_{\mathrm{St}}^h \boldsymbol{u}\right\|_{L^2(\Omega)} \leq \|\nabla \boldsymbol{u}\|_{L^2(\Omega)} + \|p\|_{L^2(\Omega)}, \tag{4.55}$$

$$\left\|I_{\mathrm{St}}^h p\right\|_{L^2(\Omega)} \leq \frac{2}{\beta_{\mathrm{is}}^h}\left(\|\nabla \boldsymbol{u}\|_{L^2(\Omega)} + \|p\|_{L^2(\Omega)}\right). \tag{4.56}$$

Proof The proof starts by using the projection as test functions and it follows then the lines of the proof of Lemma 4.7. ■

Lemma 4.43 (Approximation Properties of the Stokes Projection) *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, be a bounded domain with polyhedral and Lipschitz continuous boundary and let $(\boldsymbol{u}, p) \in V \times Q$. Assume that (4.54) is discretized with inf-sup stable conforming finite element spaces $V^h \times Q^h$, then it holds*

$$\begin{aligned}\left\|\nabla\left(\boldsymbol{u} - I_{\mathrm{St}}^h \boldsymbol{u}\right)\right\|_{L^2(\Omega)} &\leq 2\left(1 + \frac{1}{\beta_{\mathrm{is}}^h}\right) \inf_{\boldsymbol{v}^h \in V^h} \left\|\nabla(\boldsymbol{u} - \boldsymbol{v}^h)\right\|_{L^2(\Omega)} \\ &\quad + \inf_{q^h \in Q^h} \left\|p - q^h\right\|_{L^2(\Omega)},\end{aligned} \tag{4.57}$$

$$\begin{aligned}\left\|p - I_{\mathrm{St}}^h p\right\|_{L^2(\Omega)} &\leq \frac{2}{\beta_{\mathrm{is}}^h}\left[\left(1 + \frac{1}{\beta_{\mathrm{is}}^h}\right) \inf_{\boldsymbol{v}^h \in V^h} \left\|\nabla(\boldsymbol{u} - \boldsymbol{v}^h)\right\|_{L^2(\Omega)}\right. \\ &\quad \left. + \inf_{q^h \in Q^h} \left\|p - q^h\right\|_{L^2(\Omega)}\right].\end{aligned} \tag{4.58}$$

Proof Reducing (4.54) to the subspace of discretely divergence-free functions gives (4.21) with $\boldsymbol{u}^h$ replaced by $I_{\mathrm{St}}^h \boldsymbol{u}$. Now, the proof of the estimate for the velocity continuous in the same way as the proof of Theorem 4.21. The proof of the estimate for the approximation of the pressure follows the proof of Theorem 4.25. In both cases, finally the estimate (3.65) of the best approximation error of V_{div}^h is applied. ■

Remark 4.44 (Time Derivatives of the Stokes Projection and Stokes Projection of the Time Derivatives) Let $(\boldsymbol{u}, p)$ be functions depending on time which are sufficiently smooth with respect to time. Using $(\partial_t \boldsymbol{u}, \partial_t p)$ as right-hand side in the Stokes projection (4.54), one obtains the projection $\left(I_{\mathrm{St}}^h \partial_t \boldsymbol{u}, I_{\mathrm{St}}^h \partial_t p\right)$. Differentiating the definition of the Stokes projection with respect to time, interchanging differentiation in time and integration in space, and interchanging differentiation in time and in space, yields

$$a\left(\partial_t I_{\mathrm{St}}^h \boldsymbol{u}, \boldsymbol{v}^h\right) + b\left(\boldsymbol{v}^h, \partial_t I_{\mathrm{St}}^h p\right) - b\left(I_{\partial_t \mathrm{St}}^h \boldsymbol{u}, q^h\right) = a\left(\partial_t \boldsymbol{u}, \boldsymbol{v}^h\right) + b\left(\boldsymbol{v}^h, \partial_t p\right).$$

Thus, $\left(\partial_t I_{\mathrm{St}}^h \boldsymbol{u}, \partial_t I_{\mathrm{St}}^h p\right) \in V^h \times Q^h$ is also the solution of the Stokes projection with right-hand side $(\partial_t \boldsymbol{u}, \partial_t p)$. Since the Stokes projection is unique, it follows that $\left(I_{\mathrm{St}}^h \partial_t \boldsymbol{u}, I_{\mathrm{St}}^h \partial_t p\right) = \left(\partial_t I_{\mathrm{St}}^h \boldsymbol{u}, \partial_t I_{\mathrm{St}}^h p\right)$.

The commutation of projection in space and differentiation in time can be proved since all terms in the definition of the Stokes projection are linear with respect to time. It can be extended in the same way to higher order temporal derivatives. □

4.2.3 Lowest Order Non-conforming Inf-Sup Stable Pairs of Finite Element Spaces

Remark 4.45 (Non-conforming Inf-Sup Stable Finite Element Spaces of Lowest Order) Lowest order non-conforming finite element spaces are the Crouzeix–Raviart finite element P_1^{nc}/P_0 on simplicial meshes, and the Rannacher–Turek element Q_1^{rot}/Q_0 on quadrilateral and hexahedral meshes, see Sect. 3.6.5. These pairs are certainly the most popular non-conforming pairs of finite element spaces used in the simulation of incompressible flows. The fulfillment of the discrete inf-sup condition (3.51) for these pairs of spaces was shown or discussed, respectively, in Sect. 3.6.5.

This section presents exemplary the finite element error analysis for the P_1^{nc}/P_0 pair of spaces. Since $V^h \not\subset V$, properties of V are not inherited by V^h, which causes a number of technical issues in the analysis. In particular, the measure of 'V^h being not a subset of V', a so-called consistency error, has to be estimated. The proof uses that the space of conforming piecewise linear finite elements P_1 is a subspace of P_1^{nc}. □

Remark 4.46 (The Finite Element Problem) Since for non-conforming discretizations it is $V^h \not\subset V$, one has to use a definition of the finite element Stokes problem that uses mesh cell by mesh cell definitions of the bilinear forms. Let $\boldsymbol{f} \in L^2(\Omega)$, then one considers the following problem: Find $\left(\boldsymbol{u}^h, q^h\right) \in V^h \times Q^h = P_1^{\mathrm{nc}} \times P_0$ such that

$$\begin{aligned} a^h\left(\boldsymbol{u}^h, \boldsymbol{v}^h\right) + b^h\left(\boldsymbol{v}^h, p^h\right) &= \left(\boldsymbol{f}, \boldsymbol{v}^h\right) \ \forall\ \boldsymbol{v}^h \in V^h, \\ b^h\left(\boldsymbol{u}^h, q^h\right) &= 0 \qquad\quad \forall\ q^h \in Q^h, \end{aligned} \tag{4.59}$$

with the bilinear forms given in (4.18). □

Corollary 4.47 (Unique Solvability of the Finite Element Problem) *There is a unique solution $\left(\boldsymbol{u}^h, q^h\right) \in V^h \times Q^h$ of the finite element problem (4.59).*

Proof Since the discrete inf-sup condition (3.51) is satisfied, see Theorem 3.151, it is sufficient to show, according to Theorem 4.14, that $\left(a^h(\cdot,\cdot)\right)^{1/2}$ defines a norm in V^h. This property was already established in Lemma 3.150. ■

Remark 4.48 (Reduction of the Finite Element Problem to the Discretely Divergence-free Space) As for conforming finite element methods, the error analysis starts by reducing the saddle point problem (4.59) to a scalar problem in V_{div}^h: Find $\boldsymbol{u}^h \in V_{\mathrm{div}}^h$ such that

$$a^h\left(\boldsymbol{u}^h, \boldsymbol{v}^h\right) = \left(\boldsymbol{f}, \boldsymbol{v}^h\right) \quad \forall\ \boldsymbol{v}^h \in V_{\mathrm{div}}^h. \tag{4.60}$$

One obtains the same finite element velocity solution when solving problems (4.59) or (4.60). □

Remark 4.49 (Poincaré Inequality) For functions $\boldsymbol{v}^h \in V^h = P_1^{\mathrm{nc}}$ a Poincaré inequality or Poincaré–Friedrichs inequality

$$\left\|\boldsymbol{v}^h\right\|_{L^2(\Omega)} \le C \left\|\boldsymbol{v}^h\right\|_{V^h} \tag{4.61}$$

with

$$\left\|\boldsymbol{v}^h\right\|_{V^h} = \left(a^h\left(\boldsymbol{v}^h, \boldsymbol{v}^h\right)\right)^{1/2} = \left(\sum_{K\in\mathcal{T}^h} \left\|\nabla \boldsymbol{v}^h\right\|_{L^2(K)}^2\right)^{1/2}$$

holds on regular families of triangulations $\{\mathcal{T}^h\}$, e.g., see Knobloch (2001), Brenner (2003). □

Lemma 4.50 (Stability of the Finite Element Solution) *Let* $\boldsymbol{f} \in L^2(\Omega)$*, then the solution of* (4.59) *satisfies*

$$\left\|\boldsymbol{u}^h\right\|_{V^h} \le C\left\|\boldsymbol{f}\right\|_{L^2(\Omega)}, \quad \left\|p^h\right\|_{L^2(\Omega)} \le \frac{2C}{\beta_{\mathrm{is}}^h}\left\|\boldsymbol{f}\right\|_{L^2(\Omega)},$$

where C is the constant from the Poincaré inequality (4.61).

Proof Using the solution of (4.59) as test function, applying the Cauchy–Schwarz inequality (A.10), the Poincaré inequality (4.61), and Young's inequality (A.5) gives

$$\left\|\boldsymbol{u}^h\right\|_{V^h}^2 = a^h\left(\boldsymbol{u}^h, \boldsymbol{u}^h\right) = \left(\boldsymbol{f}, \boldsymbol{u}^h\right) \le \left\|\boldsymbol{f}\right\|_{L^2(\Omega)}\left\|\boldsymbol{u}^h\right\|_{L^2(\Omega)} \le C\left\|\boldsymbol{f}\right\|_{L^2(\Omega)}\left\|\boldsymbol{u}^h\right\|_{V^h},$$

which proves the velocity estimate. The estimate for the pressure is obtained with the discrete inf-sup condition (3.51), (4.59), the Cauchy–Schwarz inequality, the Poincaré inequality (4.61), and the stability estimate for the discrete velocity

$$\begin{aligned}
\left\|p^h\right\|_{L^2(\Omega)} &\le \frac{1}{\beta_{\mathrm{is}}^h}\sup_{\boldsymbol{v}^h\in V^h\setminus\{\boldsymbol{0}\}}\frac{b^h\left(\boldsymbol{v}^h, q^h\right)}{\left\|\boldsymbol{v}^h\right\|_{V^h}} = \frac{1}{\beta_{\mathrm{is}}^h}\sup_{\boldsymbol{v}^h\in V^h\setminus\{\boldsymbol{0}\}}\frac{\left(\boldsymbol{f}, \boldsymbol{v}^h\right) - a^h\left(\boldsymbol{u}^h, \boldsymbol{v}^h\right)}{\left\|\boldsymbol{v}^h\right\|_{V^h}} \\
&\le \frac{1}{\beta_{\mathrm{is}}^h}\sup_{\boldsymbol{v}^h\in V^h\setminus\{\boldsymbol{0}\}}\frac{C\left\|\boldsymbol{f}\right\|_{L^2(\Omega)}\left\|\boldsymbol{v}^h\right\|_{V^h} + \left\|\boldsymbol{u}^h\right\|_{V^h}\left\|\boldsymbol{v}^h\right\|_{V^h}}{\left\|\boldsymbol{v}^h\right\|_{V^h}} \\
&\le \frac{1}{\beta_{\mathrm{is}}^h}\left(C\left\|\boldsymbol{f}\right\|_{L^2(\Omega)} + \left\|\boldsymbol{u}^h\right\|_{V^h}\right) \le \frac{1}{\beta_{\mathrm{is}}^h}\left(2C\left\|\boldsymbol{f}\right\|_{L^2(\Omega)}\right).
\end{aligned}$$

■

Lemma 4.51 (Abstract Error Estimate, Second Lemma of Strang) *Let $\boldsymbol{u} \in V$ be the solution of* (4.4) *and $\boldsymbol{u}^h \in V^h$ be the solution of* (4.60)*. Then, there holds the error estimate*

$$\left\|\boldsymbol{u} - \boldsymbol{u}^h\right\|_{V^h} \le 2 \inf_{\boldsymbol{v}^h \in V^h_{\text{div}}} \left\|\boldsymbol{u} - \boldsymbol{v}^h\right\|_{V^h} + \inf_{\boldsymbol{v}^h \in V^h_{\text{div}}, \|\boldsymbol{v}^h\|_{V^h}=1} \left|a^h\left(\boldsymbol{u}, \boldsymbol{v}^h\right) - \left(\boldsymbol{f}, \boldsymbol{v}^h\right)\right|. \tag{4.62}$$

Proof Let $\boldsymbol{v}^h \in V^h_{\text{div}}$ be arbitrary and use the decomposition

$$\boldsymbol{u} - \boldsymbol{u}^h = \boldsymbol{u} - \boldsymbol{v}^h - \left(\boldsymbol{u}^h - \boldsymbol{v}^h\right) = \boldsymbol{u} - \boldsymbol{v}^h - \boldsymbol{\phi}^h,$$

with $\boldsymbol{\phi}^h \in V^h_{\text{div}}$. Then, one obtains, using (4.60) and the Cauchy–Schwarz inequality (A.10)

$$\begin{aligned}
\left\|\boldsymbol{\phi}^h\right\|^2_{V^h} &= a^h\left(\boldsymbol{\phi}^h, \boldsymbol{\phi}^h\right) = a^h\left(\boldsymbol{u}^h - \boldsymbol{v}^h, \boldsymbol{\phi}^h\right) \\
&= a^h\left(\boldsymbol{u} - \boldsymbol{v}^h, \boldsymbol{\phi}^h\right) + a^h\left(\boldsymbol{u}^h, \boldsymbol{\phi}^h\right) - a^h\left(\boldsymbol{u}, \boldsymbol{\phi}^h\right) \\
&= a^h\left(\boldsymbol{u} - \boldsymbol{v}^h, \boldsymbol{\phi}^h\right) + \left(\boldsymbol{f}, \boldsymbol{\phi}^h\right) - a^h\left(\boldsymbol{u}, \boldsymbol{\phi}^h\right) \\
&\le \left\|\boldsymbol{u} - \boldsymbol{v}^h\right\|_{V^h} \left\|\boldsymbol{\phi}^h\right\|_{V^h} + \left|\left(\boldsymbol{f}, \boldsymbol{\phi}^h\right) - a^h\left(\boldsymbol{u}, \boldsymbol{\phi}^h\right)\right|.
\end{aligned}$$

Dividing this estimate by $\left\|\boldsymbol{\phi}^h\right\|_{V^h} \ne 0$, using the triangle inequality

$$\left\|\boldsymbol{u} - \boldsymbol{u}^h\right\|_{V^h} \le \left\|\boldsymbol{u} - \boldsymbol{v}^h\right\|_{V^h} + \left\|\boldsymbol{u}^h - \boldsymbol{v}^h\right\|_{V^h},$$

and observing that with arbitrary $\boldsymbol{v}^h$ also $\boldsymbol{\phi}^h$ is arbitrary, gives estimate (4.62). For $\left\|\boldsymbol{\phi}^h\right\|_{V^h} = 0$, estimate (4.62) follows directly from the decomposition of the error. ■

Remark 4.52 (On the Error Estimate (4.62)*)* Error estimate (4.62) can be considered as the non-conforming analog of the Lemma of Cea, see Lemma B.12. The right-hand side consists of two parts that have to be estimated. The first term is the best approximation error in V^h_{div} and the second term is the consistency error. This term describes just the error which is committed by using a non-conforming test function in the continuous problem (4.4). □

Lemma 4.53 (Best Approximation Error Estimate for V^h_{div}) *Let $\{\mathcal{T}^h\}$ be a quasi-uniform family of triangulations and let $\boldsymbol{u} \in H^2(\Omega)$. Then, the best approximation error for V^h_{div} can be estimated in the form*

$$\inf_{\boldsymbol{v}^h \in V^h_{\text{div}}} \left\|\boldsymbol{u} - \boldsymbol{v}^h\right\|_{V^h} \le Ch\,|\boldsymbol{u}|_{H^2(\Omega)}. \tag{4.63}$$

Proof Since $\boldsymbol{u} \in V_{\text{div}}$, it follows from (3.144) that $P_E^h \boldsymbol{u} \in V_{\text{div}}^h$ since

$$0 = b\left(\boldsymbol{u}, q^h\right) = b^h\left(P_E^h \boldsymbol{u}, q^h\right) \quad \forall\, q^h \in Q^h.$$

Let $I^h : V \cap H^2(\Omega) \to P_1 \subset V^h$ denote the Lagrangian interpolation operator to the space of continuous, piecewise linear functions. Since a function from P_1 is linear on each face, the integral on the face can be computed exactly by applying the midpoint (barycenter) rule. Hence, one obtains

$$\int_E I^h \boldsymbol{u}(\boldsymbol{s})\, d\boldsymbol{s} = |E|\, I^h \boldsymbol{u}\left(\boldsymbol{m}_E\right),$$

where $\boldsymbol{m}_E$ is the barycenter of E. It follows with (3.143) that $P_E^h\left(I^h \boldsymbol{u}\right) = I^h \boldsymbol{u}$, since the values of both functions coincide at the barycenters. Using the properties derived so far, the triangle inequality, the linearity of P_E^h, estimate (3.145), and the interpolation error estimate (C.14) for I^h yields

$$\begin{aligned}
\inf_{\boldsymbol{v}^h \in V_{\text{div}}^h} \left\|\boldsymbol{u} - \boldsymbol{v}^h\right\|_{V^h} &\le \left\|\boldsymbol{u} - P_E^h \boldsymbol{u}\right\|_{V^h} \\
&\le \left\|\boldsymbol{u} - I^h \boldsymbol{u}\right\|_{V^h} + \left\|I^h \boldsymbol{u} - P_E^h \boldsymbol{u}\right\|_{V^h} \\
&= \left\|\boldsymbol{u} - I^h \boldsymbol{u}\right\|_{V^h} + \left\|P_E^h\left(I^h \boldsymbol{u}\right) - P_E^h \boldsymbol{u}\right\|_{V^h} \\
&= \left\|\boldsymbol{u} - I^h \boldsymbol{u}\right\|_{V^h} + \left\|P_E^h\left(I^h \boldsymbol{u} - \boldsymbol{u}\right)\right\|_{V^h} \\
&\le \left\|\boldsymbol{u} - I^h \boldsymbol{u}\right\|_{V^h} + \left\|I^h \boldsymbol{u} - \boldsymbol{u}\right\|_{V^h} \\
&\le Ch\, |\boldsymbol{u}|_{H^2(\Omega)}\,.
\end{aligned}$$

■

Remark 4.54 (On Estimate (4.63)) Note that the constant on the right-hand side of estimate (4.63) does not depend on the discrete inf-sup constant. This property is in contrast to the general best approximation error estimate (3.65) for conforming finite element spaces. Recall, it was shown in Theorem 3.151 that the discrete inf-sup constant for the Crouzeix–Raviart pair of finite element spaces anyway does not depend on the concrete triangulation since it can be chosen to be the continuous inf-sup constant. □

Lemma 4.55 (Consistency Error Estimate) *Let $(\boldsymbol{u}, p)$ be a sufficiently smooth solution of the Stokes equations (4.2) with $\boldsymbol{u} \in C^1\left(\overline{\Omega}\right) \cap V$, $p \in C\left(\overline{\Omega}\right) \cap Q$. Consider a family of quasi-uniform triangulations, then it holds for all $\boldsymbol{v} \in V_{\text{div}} \oplus V_{\text{div}}^h$*

$$\left|a^h\left(\boldsymbol{u}, \boldsymbol{v}\right) - \left(\boldsymbol{f}, \boldsymbol{v}\right)\right| \le Ch\left(|\boldsymbol{u}|_{H^2(\Omega)} + \|\nabla p\|_{L^2(\Omega)}\right) \|\boldsymbol{v}\|_{V^h}\,, \tag{4.64}$$

where the constant depends only on the regularity of the family of triangulations.

Proof Let $\boldsymbol{v} \in V_{\text{div}} \oplus V_{\text{div}}^h$ be arbitrary, then it is, using the momentum equation of the Stokes problem (4.1)

$$\begin{aligned} a^h(\boldsymbol{u}, \boldsymbol{v}) - (\boldsymbol{f}, \boldsymbol{v}) &= \sum_{K \in \mathcal{T}^h} (\nabla \boldsymbol{u}, \nabla \boldsymbol{v})_K - (\boldsymbol{f}, \boldsymbol{v})_K \\ &= \sum_{K \in \mathcal{T}^h} (\nabla \boldsymbol{u}, \nabla \boldsymbol{v})_K - (-\Delta \boldsymbol{u} + \nabla p, \boldsymbol{v})_K . \end{aligned} \tag{4.65}$$

Applying integration by parts, mesh cell by mesh cell, yields

$$\begin{aligned} &\sum_{K \in \mathcal{T}^h} (\nabla \boldsymbol{u}, \nabla \boldsymbol{v})_K + (\Delta \boldsymbol{u}, \boldsymbol{v})_K \\ &= \sum_{E \in \mathcal{E}^h} \int_E [|(\nabla \boldsymbol{u} \boldsymbol{n}_E) \cdot \boldsymbol{v}|]_E \, d\boldsymbol{s} + \sum_{E \in \overline{\mathcal{E}}^h \setminus \mathcal{E}^h} \int_E (\nabla \boldsymbol{u} \boldsymbol{n}_E) \cdot \boldsymbol{v} \, d\boldsymbol{s}, \end{aligned} \tag{4.66}$$

where the unit normals $\boldsymbol{n}_E$ for $E \in \mathcal{E}^h$ were chosen arbitrarily but fixed and the normals for $E \in \overline{\mathcal{E}}^h \setminus \mathcal{E}^h$ are the outward pointing unit normals. Note that changing the normal for an interior face changes the both, the sign of the normal and the sign of the jump, such that one obtains the same result as with the other normal.

Let

$$\overline{\boldsymbol{v}}_E = \frac{1}{|E|} \int_E \boldsymbol{v}(\boldsymbol{s}) \, d\boldsymbol{s}, \quad E \in \mathcal{E}^h,$$

be the integral mean value of $\boldsymbol{v}$ on E. Note that the integral mean value is well defined for functions on V^h, since it is the nodal functional for defining the space P_1^{nc} and this functional has to be continuous. It follows that

$$\int_E (\boldsymbol{v} - \overline{\boldsymbol{v}}_E)(\boldsymbol{s}) \, d\boldsymbol{s} = 0. \tag{4.67}$$

In addition, let $I^h : V \cap H^2(\Omega) \to P_1 \subset V^h$ be the Lagrangian interpolation operator to the space of continuous, piecewise linear functions. Then, for each mesh cell and each face E of the mesh cell, $\left(\nabla I^h \boldsymbol{u}\right) \boldsymbol{n}_E$ is constant. For (4.66), it follows that

$$\begin{aligned} &\sum_{K \in \mathcal{T}^h} (\nabla \boldsymbol{u}, \nabla \boldsymbol{v})_K + (\Delta \boldsymbol{u}, \boldsymbol{v})_K \\ &= \sum_{E \in \mathcal{E}^h} \int_E \left[\left|\nabla \left(\boldsymbol{u} - I^h \boldsymbol{u}\right) \boldsymbol{n}_E \cdot (\boldsymbol{v} - \overline{\boldsymbol{v}}_E)\right|\right]_E (\boldsymbol{s}) \, d\boldsymbol{s} \\ &\quad + \sum_{E \in \overline{\mathcal{E}}^h \setminus \mathcal{E}^h} \int_E \nabla \left(\boldsymbol{u} - I^h \boldsymbol{u}\right) \boldsymbol{n}_E \cdot (\boldsymbol{v} - \overline{\boldsymbol{v}}_E)(\boldsymbol{s}) \, d\boldsymbol{s}, \end{aligned} \tag{4.68}$$

since for the integral on the boundary faces, one has with (4.67) and $\overline{\boldsymbol{v}}_E = \mathbf{0}$

$$\begin{aligned}
&\int_E \nabla\left(\boldsymbol{u} - I^h\boldsymbol{u}\right)\boldsymbol{n}_E \cdot (\boldsymbol{v} - \overline{\boldsymbol{v}}_E)(\boldsymbol{s})\, d\boldsymbol{s} \\
&= \int_E (\nabla \boldsymbol{u}\boldsymbol{n}_E)\cdot \boldsymbol{v}(\boldsymbol{s})\, d\boldsymbol{s} - \int_E (\nabla \boldsymbol{u}\boldsymbol{n}_E)\cdot \overline{\boldsymbol{v}}_E(\boldsymbol{s})\, d\boldsymbol{s} - \nabla I^h \boldsymbol{u}\boldsymbol{n}_E \cdot \int_E (\boldsymbol{v} - \overline{\boldsymbol{v}}_E)\ d\boldsymbol{s} \\
&= \int_E (\nabla \boldsymbol{u}\boldsymbol{n}_E)\cdot \boldsymbol{v}(\boldsymbol{s})\, d\boldsymbol{s}.
\end{aligned}$$

For the interior edges, one obtains with (4.67)

$$\begin{aligned}
&\int_E \left[\left|\nabla\left(\boldsymbol{u} - I^h\boldsymbol{u}\right)\boldsymbol{n}_E \cdot (\boldsymbol{v} - \overline{\boldsymbol{v}}_E)\right|\right]_E(\boldsymbol{s})\, d\boldsymbol{s} \\
&= \int_E [|\nabla \boldsymbol{u}\boldsymbol{n}_E \cdot \boldsymbol{v}|]_E(\boldsymbol{s})\, d\boldsymbol{s} - \int_E [|\nabla \boldsymbol{u}\boldsymbol{n}_E \cdot \overline{\boldsymbol{v}}_E|]_E(\boldsymbol{s})\, d\boldsymbol{s} \\
&\quad - \int_E \left[\left|\nabla\left(I^h\boldsymbol{u}\right)\boldsymbol{n}_E \cdot (\boldsymbol{v} - \overline{\boldsymbol{v}}_E)\right|\right]_E(\boldsymbol{s})\, d\boldsymbol{s} \\
&= \int_E [|\nabla \boldsymbol{u}\boldsymbol{n}_E \cdot \boldsymbol{v}|]_E\ d\boldsymbol{s} - \int_E [|\nabla \boldsymbol{u}\boldsymbol{n}_E \cdot \overline{\boldsymbol{v}}_E|]_E(\boldsymbol{s})\, d\boldsymbol{s} \\
&\quad - \nabla\left(I^h\boldsymbol{u}\right)\boldsymbol{n}_E\big|_{K_1} \int_E (\boldsymbol{v} - \overline{\boldsymbol{v}}_E)(\boldsymbol{s})\, d\boldsymbol{s} + \nabla\left(I^h\boldsymbol{u}\right)\boldsymbol{n}_E\big|_{K_2} \int_E (\boldsymbol{v} - \overline{\boldsymbol{v}}_E)(\boldsymbol{s})\, d\boldsymbol{s} \\
&= \int_E [|\nabla \boldsymbol{u}\boldsymbol{n}_E \cdot \boldsymbol{v}|]_E(\boldsymbol{s})\, d\boldsymbol{s}.
\end{aligned}$$

The second term in the next to last equation vanishes because the function in the jump is continuous almost everywhere, thus the jump is zero almost everywhere, and the last two terms vanish because of (4.67). Both terms on the right-hand side of (4.68) can be estimated by

$$Ch_E \left\|\nabla \boldsymbol{v}\right\|_{L^2(K)} |\boldsymbol{u}|_{H^2(K)}, \quad E \subset K,$$

see Lemma 4.58 below, such that one can estimate (4.66) with

$$\sum_{K\in\mathcal{T}^h} (\nabla \boldsymbol{u}, \nabla \boldsymbol{v})_K + (\Delta \boldsymbol{u}, \boldsymbol{v})_K \le Ch\, |\boldsymbol{u}|_{H^2(\Omega)} \left\|\boldsymbol{v}\right\|_{V^h}. \tag{4.69}$$

Consider now the last term on the right-hand side of (4.65). Integration by parts yields

$$\begin{aligned}
\sum_{K\in\mathcal{T}^h} (\nabla p, \boldsymbol{v})_K = \sum_{E\in\mathcal{E}^h} \int_E [|p\boldsymbol{v}\cdot\boldsymbol{n}_E|]_E\ ds + \sum_{E\in\overline{\mathcal{E}}^h\setminus\mathcal{E}^h} \int_E p\boldsymbol{v}\cdot\boldsymbol{n}_E\, ds \\
- \sum_{K\in\mathcal{T}^h} (\nabla\cdot\boldsymbol{v}, p)_K\,.
\end{aligned} \tag{4.70}$$

Let $P_{L^2}^h : Q \to Q^h$ be the $L^2(\Omega)$ projection of pressure functions to the piecewise constant finite element pressure. With the similar arguments as for estimating the integrals for the velocity, see Lemma 4.58 below, one obtains

$$
\begin{aligned}
&\sum_{E\in\mathcal{E}^h} \int_E [|p\boldsymbol{v}\cdot\boldsymbol{n}_E|]_E \, ds + \sum_{E\in\overline{\mathcal{E}}^h\setminus\mathcal{E}^h} \int_E p\boldsymbol{v}\cdot\boldsymbol{n}_E \, ds \\
&= \sum_{E\in\mathcal{E}^h} \int_E \left[\left|\left(p - P_{L^2}^h p\right)(\boldsymbol{v} - \overline{\boldsymbol{v}}_E)\cdot\boldsymbol{n}_E\right|\right]_E ds \\
&\quad + \sum_{E\in\overline{\mathcal{E}}^h\setminus\mathcal{E}^h} \int_E \left(p - P_{L^2}^h p\right)(\boldsymbol{v} - \overline{\boldsymbol{v}}_E)\cdot\boldsymbol{n}_E \, ds \\
&\le Ch \|\nabla p\|_{L^2(\Omega)} \|\boldsymbol{v}\|_{V^h}.
\end{aligned} \tag{4.71}
$$

If $\boldsymbol{v} \in V_{\mathrm{div}}$, then the last term on the right-hand side of (4.70) vanishes by the definition of this space. Thus, let $\boldsymbol{v} \in V_{\mathrm{div}}^h$, then one obtains, using the definition of V_{div}^h, the Cauchy–Schwarz inequality (A.10), the Cauchy–Schwarz inequality for sums (A.2), estimate (3.40), and the estimate for the $L^2(\Omega)$ projection (C.28)

$$
\begin{aligned}
\sum_{K\in\mathcal{T}^h} (\nabla\cdot\boldsymbol{v}, p)_K &= \sum_{K\in\mathcal{T}^h} \left(\nabla\cdot\boldsymbol{v}, p - P_{L^2}^h p\right)_K \\
&\le \sum_{K\in\mathcal{T}^h} \|\nabla\cdot\boldsymbol{v}\|_{L^2(K)} \left\|p - P_{L^2}^h p\right\|_{L^2(K)} \\
&\le \left\|p - P_{L^2}^h p\right\|_{L^2(\Omega)} \left(\sum_{K\in\mathcal{T}^h} \|\nabla\cdot\boldsymbol{v}\|_{L^2(K)}^2\right)^{1/2} \\
&\le Ch \|\nabla p\|_{L^2(\Omega)} \|\boldsymbol{v}\|_{V^h}.
\end{aligned} \tag{4.72}
$$

Collecting estimates (4.69), (4.71), and (4.72) gives the statement of the lemma. ■

Remark 4.56 (On the Smoothness Assumption in Lemma 4.55) From the smoothness assumptions in the formulation of Lemma 4.55, it follows that there are no jumps of the gradient of the velocity and the pressure across faces. From the Sobolev imbedding Theorem A.42 iii), it follows that these conditions are satisfied if $\boldsymbol{u} \in H^{d/2+1+\varepsilon}(\Omega)$ and $p \in H^{d/2+\varepsilon}(\Omega)$.

Strictly speaking, the assumption on continuity can be relax somewhat. In the proof of Lemma 4.55 it is sufficient that the gradient of the velocity and the pressure do not possess jump discontinuities. E.g., in the two-dimensional case, this property is given if $\boldsymbol{u} \in H^2(\Omega)$ and $p \in H^1(\Omega)$. □

Remark 4.57 (On Test Functions from V^h in Lemma 4.55) The property of $\boldsymbol{v}$ being (discretely) divergence-free was used only in estimating the last term of (4.70). Estimates (4.69) and (4.71) can be derived in the same way for functions $\boldsymbol{v} \in V^h$. □

Lemma 4.58 (Estimate of the Jump Terms in the Proof of Lemma 4.55) *Let the assumptions and notations like in Lemma 4.55, then*

$$\int_E \left[\left|\nabla \left(\boldsymbol{u} - I^h\boldsymbol{u}\right) \boldsymbol{n}_E \cdot (\boldsymbol{v} - \overline{\boldsymbol{v}}_E)\right|\right]_E (\boldsymbol{s})\, d\boldsymbol{s} \le C h_E \left\|\nabla \boldsymbol{v}\right\|_{L^2(K)} \left|\boldsymbol{u}\right|_{H^2(K)}$$

for $E \in \overline{\mathcal{E}}$ and $E \subset \partial K$.

Proof Using the Cauchy–Schwarz inequality (A.10) and $\|\boldsymbol{n}_E\|_2 = 1$ gives

$$\int_E \left[\left|\nabla \left(\boldsymbol{u} - I^h\boldsymbol{u}\right) \boldsymbol{n}_E \cdot (\boldsymbol{v} - \overline{\boldsymbol{v}}_E)\right|\right]_E (\boldsymbol{s})\, d\boldsymbol{s} \le \left\|\nabla \left(\boldsymbol{u} - I^h\boldsymbol{u}\right)\right\|_{L^2(E)} \left\|\boldsymbol{v} - \overline{\boldsymbol{v}}_E\right\|_{L^2(E)} . \tag{4.73}$$

Let $F_K : \hat{K} \to K$ be an affine transform which maps:

- in 2d: the edge $\hat{E}$ with the vertices $(0,0)^T$ and $(1,0)^T$ of the standard reference simplex to E,
- in 3d: the face $\hat{E}$ with the vertices $(0,0,0)^T$, $(1,0,0)^T$, and $(0,1,0)^T$ of the standard reference tetrahedron to E.

Setting $\hat{\boldsymbol{w}} = \boldsymbol{w} \circ F_K$, one can apply the standard transform of integrals (B.32) to the edge or to the face, respectively, to obtain

$$\overline{\boldsymbol{w}}_E = \frac{1}{|E|} \int_E \boldsymbol{w}(\boldsymbol{s})\, d\boldsymbol{s} = \frac{1}{\left|\hat{E}\right|} \int_{\hat{E}} \hat{\boldsymbol{w}}(\hat{\boldsymbol{s}})\, d\hat{\boldsymbol{s}},$$

where (B.36) was used. Since $\overline{\boldsymbol{w}}_E \in \mathbb{R}$, one gets

$$\int_{\hat{E}} \left(\hat{\boldsymbol{w}} - \overline{\boldsymbol{w}}_E\right)(\hat{\boldsymbol{s}})\, d\hat{\boldsymbol{s}} = \int_{\hat{E}} \hat{\boldsymbol{w}}(\hat{\boldsymbol{s}})\, d\hat{\boldsymbol{s}} - \overline{\boldsymbol{w}}_E \left|\hat{E}\right| = 0.$$

From Lemma C.3, it follows that there is a constant $\hat{C}$ such that

$$\left\|\hat{\phi}\right\|_{L^2(K)} \le \hat{C} \left\|\nabla \hat{\phi}\right\|_{L^2(K)} , \tag{4.74}$$

for all $\hat{\phi} \in H^1(\hat{K})$ with $\int_{\hat{E}} \hat{\phi}(\hat{s})\, d\hat{s} = 0$.

Starting the estimate of the second term on the right-hand side of (4.73) with the transform to reference configuration, applying (C.11), the trace theorem, Theorem A.34, utilizing (4.74), using that $\overline{\boldsymbol{v}}_E$ is a constant, applying the back transform, and using (C.12) together with $h_K \le h_E$, which follows from the quasi-uniformity of the family of triangulations, gives

$$\begin{aligned}
\|\boldsymbol{v} - \overline{\boldsymbol{v}}_E\|_{L^2(E)} &\le h_E^{d/2-1/2} \|\hat{\boldsymbol{v}} - \overline{\boldsymbol{v}}_E\|_{L^2(\hat{E})} \\
&\le Ch_E^{d/2-1/2} \left(\|\hat{\boldsymbol{v}} - \overline{\boldsymbol{v}}_E\|^2_{L^2(\hat{K})} + \|\nabla(\hat{\boldsymbol{v}} - \overline{\boldsymbol{v}}_E)\|^2_{L^2(\hat{K})}\right)^{d/2-1/2} \\
&\le C\left(1+\hat{C}\right) h_E^{d/2-1/2} \|\nabla(\hat{\boldsymbol{v}} - \overline{\boldsymbol{v}}_E)\|_{L^2(\hat{K})} \\
&= C\left(1+\hat{C}\right) h_E^{d/2-1/2} \|\nabla\hat{\boldsymbol{v}}\|_{L^2(\hat{K})} \\
&\le Ch_E^{d/2-1/2} h_E^{1-d/2} \|\nabla\boldsymbol{v}\|_{L^2(K)} = Ch_E^{1/2}\|\nabla\boldsymbol{v}\|_{L^2(K)} .
\end{aligned}$$

Similarly, using that $\widehat{I^h\boldsymbol{u}}$ is linear on $\hat{K}$ and the interpolation estimate (C.4), one obtains for the first term on the right-hand side of (4.73)

$$\begin{aligned}
&\left\|\nabla\left(\boldsymbol{u} - I^h\boldsymbol{u}\right)\right\|_{L^2(E)} \\
&\le h_E^{d/2-3/2} \left\|\nabla\left(\hat{\boldsymbol{u}} - \widehat{I^h\boldsymbol{u}}\right)\right\|_{L^2(\hat{E})} \\
&\le Ch_E^{d/2-3/2} \left\|\hat{\boldsymbol{u}} - \widehat{I^h\boldsymbol{u}}\right\|^2_{H^2(\hat{K})} \\
&= Ch_E^{d/2-3/2} \left(\left\|\hat{\boldsymbol{u}} - \widehat{I^h\boldsymbol{u}}\right\|^2_{L^2(\hat{K})} + \left\|\nabla\left(\hat{\boldsymbol{u}} - \widehat{I^h\boldsymbol{u}}\right)\right\|^2_{L^2(\hat{K})} + |\hat{\boldsymbol{u}}|^2_{H^2(\hat{K})}\right)^{1/2} \\
&\le Ch_E^{d/2-3/2} |\hat{\boldsymbol{u}}|_{H^2(\hat{K})} \\
&\le Ch_E^{d/2-3/2} h_E^{2-d/2} |\boldsymbol{u}|_{H^2(K)} = Ch_E^{1/2} |\boldsymbol{u}|_{H^2(K)} .
\end{aligned}$$

Combining the last two estimates proves the statement of the lemma. ∎

Theorem 4.59 (Error Estimate for the V^h Norm of the Velocity) *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, be a bounded domain with polyhedral and Lipschitz continuous boundary and let $(\boldsymbol{u}, p)$, the unique solution of the Stokes problem (4.2), be sufficiently smooth in the sense of Remark 4.56. Consider a quasi-uniform family of triangulations and let $\left(\boldsymbol{u}^h, p^h\right) \in P_1^{\mathrm{nc}} \times P_0$ be the unique solution of the finite element problem (4.59), then it holds the error estimate*

$$\left\|\boldsymbol{u} - \boldsymbol{u}^h\right\|_{V^h} \le Ch\left(|\boldsymbol{u}|_{H^2(\Omega)} + \|\nabla p\|_{L^2(\Omega)}\right), \tag{4.75}$$

where the constant depends only on Ω and on the regularity of the family of triangulations.

Proof Estimate (4.75) follows by inserting (4.63) and (4.64) in (4.62). ∎

Theorem 4.60 (Error Estimate for the $L^2(\Omega)$ Norm of the Pressure) *Let the conditions of Theorem 4.59 be satisfied, then it holds*

$$\left\|p-p^h\right\|_{L^2(\Omega)} \le Ch\left(1+\frac{1}{\beta_{\mathrm{is}}^h}\right)\left(|\boldsymbol{u}|_{H^2(\Omega)}+\|\nabla p\|_{L^2(\Omega)}\right), \tag{4.76}$$

where the constant C depends only on Ω and on the regularity of the family of triangulations and for the discrete inf-sup constant it holds $\beta_{\mathrm{is}}^h = \beta_{\mathrm{is}}$, see Theorem 3.151.

Proof Let $P_{L^2}^h p$ denote the $L^2(\Omega)$ interpolant of $p \in Q$ in Q^h. Then, one obtains with the triangle inequality

$$\left\|p-p^h\right\|_{L^2(\Omega)} \le \left\|p-P_{L^2}^h p\right\|_{L^2(\Omega)} + \left\|p^h-P_{L^2}^h p\right\|_{L^2(\Omega)}.$$

Using (C.28), the first term on the right-hand side can be bounded in the form

$$\left\|p-P_{L^2}^h p\right\|_{L^2(\Omega)} \le Ch\left\|\nabla p\right\|_{L^2(\Omega)}. \tag{4.77}$$

From the discrete inf-sup condition (3.51), it follows for the second term that

$$\left\|p^h-P_{L^2}^h p\right\|_{L^2(\Omega)} \le \frac{1}{\beta_{\mathrm{is}}^h}\sup_{\boldsymbol{v}^h\in V^h\setminus\{\boldsymbol{0}\}}\frac{b^h\left(\boldsymbol{v}^h,p^h-P_{L^2}^h p\right)}{\left\|\boldsymbol{v}^h\right\|_{V^h}}. \tag{4.78}$$

The numerator will be split into two parts

$$b^h\left(\boldsymbol{v}^h,p^h-P_{L^2}^h p\right) = b^h\left(\boldsymbol{v}^h,p^h-p\right)+b^h\left(\boldsymbol{v}^h,p-P_{L^2}^h p\right),$$

which are estimated separately. One gets with the Cauchy–Schwarz inequality (A.10), with (3.40), and with (4.77)

$$\begin{aligned}\left|b^h\left(\boldsymbol{v}^h,p-P_{L^2}^h p\right)\right| &\le \sqrt{d}\left\|\boldsymbol{v}^h\right\|_{V^h}\left\|p-P_{L^2}^h p\right\|_{L^2(\Omega)}\\ &\le Ch\left\|\nabla p\right\|_{L^2(\Omega)}\left\|\boldsymbol{v}^h\right\|_{V^h}.\end{aligned} \tag{4.79}$$

For the other term, one inserts the finite element problem (4.59) and the continuous problem (4.1)

$$\begin{aligned}b^h\left(\boldsymbol{v}^h,p^h-p\right) &= -a^h\left(\boldsymbol{u}^h,\boldsymbol{v}^h\right)+\left(\boldsymbol{f},\boldsymbol{v}^h\right)-b^h\left(\boldsymbol{v}^h,p\right)\\ &= a^h\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{v}^h\right)-a^h\left(\boldsymbol{u},\boldsymbol{v}^h\right)+\left(\boldsymbol{f},\boldsymbol{v}^h\right)-b^h\left(\boldsymbol{v}^h,p\right)\\ &= a^h\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{v}^h\right)+\sum_{K\in\mathcal{T}^h}\Big[-\left(\nabla\boldsymbol{u},\nabla\boldsymbol{v}^h\right)_K\\ &\quad -\left(\Delta\boldsymbol{u},\boldsymbol{v}^h\right)_K+\left(\nabla p,\boldsymbol{v}^h\right)_K+\left(\nabla\cdot\boldsymbol{v}^h,p\right)_K\Big].\end{aligned} \tag{4.80}$$

The first term of (4.80) is bounded with the Cauchy–Schwarz inequality and error estimate (4.75)

$$
\begin{aligned}
a^h\left(\boldsymbol{u}-\boldsymbol{u}^h, \boldsymbol{v}^h\right) &\leq \left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{V^h}\left\|\boldsymbol{v}^h\right\|_{V^h} \\
&\leq C\left(|\boldsymbol{u}|_{H^2(\Omega)}+\|\nabla p\|_{L^2(\Omega)}\right)\left\|\boldsymbol{v}^h\right\|_{V^h}. \qquad (4.81)
\end{aligned}
$$

For the second term on the right-hand side of (4.80), integration by parts is applied, and the estimates (4.69) and (4.71) for $\boldsymbol{v}^h \in V^h$, see Remark 4.57, are used

$$
\begin{aligned}
&\sum_{K\in\mathcal{T}^h}\left[-\left(\nabla\boldsymbol{u},\nabla\boldsymbol{v}^h\right)_K+\left(\nabla\cdot\boldsymbol{v}^h,p\right)_K-\left(\Delta\boldsymbol{u},\boldsymbol{v}^h\right)_K+\left(\nabla p,\boldsymbol{v}^h\right)_K\right] \\
&=\sum_{E\in\mathcal{E}^h}\left[-\int_E [|\nabla\boldsymbol{u}\boldsymbol{n}_E\cdot\boldsymbol{v}^h|]_E(\boldsymbol{s})\,d\boldsymbol{s}+\int_E [|p\boldsymbol{v}^h\cdot\boldsymbol{n}_E|]_E(\boldsymbol{s})\,d\boldsymbol{s}\right] \\
&\quad+\sum_{E\in\overline{\mathcal{E}}^h\setminus\mathcal{E}^h}\left[-\int_E \nabla\boldsymbol{u}\boldsymbol{n}_E\cdot\boldsymbol{v}^h(\boldsymbol{s})\,d\boldsymbol{s}+\int_E p\boldsymbol{v}^h\cdot\boldsymbol{n}_E(\boldsymbol{s})\,d\boldsymbol{s}\right] \\
&\leq Ch\left(|\boldsymbol{u}|_{H^2(\Omega)}+\|\nabla p\|_{L^2(\Omega)}\right)\left\|\boldsymbol{v}^h\right\|_{V^h}. \qquad (4.82)
\end{aligned}
$$

Inserting (4.79)–(4.82) in (4.78) gives

$$
\begin{aligned}
\left\|p^h-P_{L^2}^h p\right\|_{L^2(\Omega)} &\leq \frac{Ch}{\beta_{\mathrm{is}}^h}\sup_{\boldsymbol{v}^h\in V^h\setminus\{\boldsymbol{0}\}}\frac{\left(|\boldsymbol{u}|_{H^2(\Omega)}+\|\nabla p\|_{L^2(\Omega)}\right)\left\|\boldsymbol{v}^h\right\|_{V^h}}{\|\boldsymbol{v}^h\|_{V^h}} \\
&=\frac{Ch}{\beta_{\mathrm{is}}^h}\left(|\boldsymbol{u}|_{H^2(\Omega)}+\|\nabla p\|_{L^2(\Omega)}\right).
\end{aligned}
$$

Together with (4.77), one obtains the statement of the theorem. ■

Theorem 4.61 (Error Estimate for the $L^2(\Omega)$ Norm of the Velocity) *Let $(\boldsymbol{\phi}_{\hat{f}}, \xi_{\hat{f}})$ be the regular solution of the dual Stokes equations (4.27), then there holds the error estimate*

$$
\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{L^2(\Omega)} \leq Ch^2\left(|\boldsymbol{u}|_{H^2(\Omega)}+\|\nabla p\|_{L^2(\Omega)}\right). \qquad (4.83)
$$

Proof Consider the weak form of the dual problem: Find $\boldsymbol{\phi}_{\hat{f}} \in V_{\mathrm{div}}$ such that

$$
a\left(\boldsymbol{v},\boldsymbol{\phi}_{\hat{f}}\right)=\left(\hat{\boldsymbol{f}},\boldsymbol{v}\right) \quad \forall\ \boldsymbol{v}\in V_{\mathrm{div}}
$$

and its Galerkin discretization: Find $\boldsymbol{\phi}_{\hat{f}}^h \in V_{\mathrm{div}}^h$ such that

$$
a^h\left(\boldsymbol{v}^h,\boldsymbol{\phi}_{\hat{f}}^h\right)=\left(\hat{\boldsymbol{f}},\boldsymbol{v}^h\right) \quad \forall\ \boldsymbol{v}^h\in V_{\mathrm{div}}^h.
$$

By definition, it is

$$\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{L^2(\Omega)}=\sup_{\hat{\boldsymbol{f}}\in L^2(\Omega)\setminus\{\boldsymbol{0}\}}\frac{\left(\hat{\boldsymbol{f}},\boldsymbol{u}-\boldsymbol{u}^h\right)}{\left\|\hat{\boldsymbol{f}}\right\|_{L^2(\Omega)}}.$$

For the numerator, one gets, using the definition of the dual problem and its discretization

$$\begin{aligned}
\left(\hat{\boldsymbol{f}},\boldsymbol{u}-\boldsymbol{u}^h\right) &= a\left(\boldsymbol{u},\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}\right)-a^h\left(\boldsymbol{u}^h,\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}^h\right)\\
&= a^h\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}-\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}^h\right)+a^h\left(\boldsymbol{u}^h,\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}-\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}^h\right)+a^h\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}^h\right)\\
&= a^h\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}-\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}^h\right)+a^h\left(\boldsymbol{u}^h,\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}\right)-a^h\left(\boldsymbol{u},\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}\right)+a^h\left(\boldsymbol{u},\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}\right)\\
&\quad -a^h\left(\boldsymbol{u}^h,\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}^h\right)+a^h\left(\boldsymbol{u},\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}^h\right)-a^h\left(\boldsymbol{u},\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}\right)+a^h\left(\boldsymbol{u},\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}\right)-a^h\left(\boldsymbol{u}^h,\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}^h\right)\\
&= a^h\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}-\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}^h\right)+\left[-a^h\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}\right)+\left(\hat{\boldsymbol{f}},\boldsymbol{u}\right)-\left(\hat{\boldsymbol{f}},\boldsymbol{u}^h\right)\right]\\
&\quad +\left[-a^h\left(\boldsymbol{u},\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}-\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}^h\right)+\left(\hat{\boldsymbol{f}},\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}\right)-\left(\hat{\boldsymbol{f}},\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}^h\right)\right].
\end{aligned}$$

Now, the application of the Cauchy–Schwarz inequality (A.10), the consistency error estimate (4.64), and the finite element error estimate (4.75) (both estimates for the primal and the dual problem) gives

$$\begin{aligned}
&\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{L^2(\Omega)}\\
&\le \sup_{\hat{\boldsymbol{f}}\in L^2(\Omega)\setminus\{\boldsymbol{0}\}}\frac{1}{\left\|\hat{\boldsymbol{f}}\right\|_{L^2(\Omega)}}\Bigg(\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{V^h}\left\|\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}-\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}^h\right\|_{V^h}\\
&\quad +\left|a^h\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}\right)-\left(\hat{\boldsymbol{f}},\boldsymbol{u}-\boldsymbol{u}^h\right)\right|+\left|a^h\left(\boldsymbol{u},\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}-\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}^h\right)-\left(\hat{\boldsymbol{f}},\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}-\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}^h\right)\right|\Bigg)\\
&\le \sup_{\hat{\boldsymbol{f}}\in L^2(\Omega)\setminus\{\boldsymbol{0}\}}\frac{1}{\left\|\hat{\boldsymbol{f}}\right\|_{L^2(\Omega)}}\Bigg(\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{V^h}\left\|\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}-\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}^h\right\|_{V^h}\\
&\quad +Ch\left(\left|\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}\right|_{H^2(\Omega)}+\left\|\nabla\xi_{\hat{\boldsymbol{f}}}\right\|_{L^2(\Omega)}\right)\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{V^h}\\
&\quad +Ch\left(|\boldsymbol{u}|_{H^2(\Omega)}+\|\nabla p\|_{L^2(\Omega)}\right)\left\|\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}-\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}^h\right\|_{V^h}\Bigg) \qquad (4.84)\\
&\le Ch^2\left(|\boldsymbol{u}|_{H^2(\Omega)}+\|\nabla p\|_{L^2(\Omega)}\right)\sup_{\hat{\boldsymbol{f}}\in L^2(\Omega)\setminus\{\boldsymbol{0}\}}\frac{1}{\left\|\hat{\boldsymbol{f}}\right\|_{L^2(\Omega)}}\left(\left|\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}\right|_{H^2(\Omega)}+\left\|\nabla\xi_{\hat{\boldsymbol{f}}}\right\|_{L^2(\Omega)}\right).
\end{aligned}$$

Since the map (4.28) is an isomorphism from $\left(H^2(\Omega)\cap V\right)\times\left(H^1(\Omega)\cap Q\right)$ onto $L^2(\Omega)$ there is a constant C with

$$\left\|\boldsymbol{\phi}_{\hat{f}}\right\|_{H^2(\Omega)}+\left\|\xi_{\hat{f}}\right\|_{H^1(\Omega)}\le C\left\|\hat{\boldsymbol{f}}\right\|_{L^2(\Omega)}.$$

Inserting this property in (4.84) proves (4.83). ■

Lemma 4.62 (Divergence-free Property of Functions From V_{div}^h) *Functions from V_{div}^h are pointwise divergence-free in the interior of mesh cells.*

Proof Let $\boldsymbol{v}^h\in V_{\mathrm{div}}^h$ be an arbitrary function and $K_1\in\mathcal{T}^h$ be an arbitrary mesh cell. In the same way as in Remark 4.32, one finds that

$$\sum_{K\in\mathcal{T}^h}\int_K\nabla\cdot\boldsymbol{v}^h(\boldsymbol{x})\,d\boldsymbol{x}=\frac{|\Omega|}{|K_1|}\int_{K_1}\nabla\cdot\boldsymbol{v}^h(\boldsymbol{x})\,d\boldsymbol{x}. \tag{4.85}$$

Using integration by parts yields

$$\begin{aligned}\sum_{K\in\mathcal{T}^h}\int_K\nabla\cdot\boldsymbol{v}^h(\boldsymbol{x})\,d\boldsymbol{x}&=\sum_{K\in\mathcal{T}^h}\sum_{E\subset\partial K}\int_E\boldsymbol{v}^h\cdot\boldsymbol{n}_E(\boldsymbol{s})\,d\boldsymbol{s}\\&=\sum_{E\in\mathcal{E}^h}\int_E\left[\left|\boldsymbol{v}^h\cdot\boldsymbol{n}_E\right|\right]_E(\boldsymbol{s})\,d\boldsymbol{s}+\sum_{E\in\overline{\mathcal{E}}^h\setminus\mathcal{E}^h}\int_E\left(\boldsymbol{v}^h\cdot\boldsymbol{n}_E\right)(\boldsymbol{s})\,d\boldsymbol{s}.\end{aligned}$$

The jumps $\left[\left|\boldsymbol{v}^h\cdot\boldsymbol{n}_E\right|\right]_E$ on the interior faces and $\left(\boldsymbol{v}^h\cdot\boldsymbol{n}_E\right)$ on boundary faces are linear functions. Thus, the integrals can be evaluated exactly with the mid point rule, leading to

$$\sum_{K\in\mathcal{T}^h}\int_K\nabla\cdot\boldsymbol{v}^h(\boldsymbol{x})\,d\boldsymbol{x}=\sum_{E\in\mathcal{E}^h}|E|\left[\left|\boldsymbol{v}^h\cdot\boldsymbol{n}_E\right|\right]_E(\boldsymbol{m}_E)+\sum_{E\in\overline{\mathcal{E}}^h\setminus\mathcal{E}^h}|E|\left(\boldsymbol{v}^h\cdot\boldsymbol{n}_E\right)(\boldsymbol{m}_E).$$

By definition of V^h, see (B.19), the functions are continuous in the barycenter of all faces and the application of the homogeneous Dirichlet boundary condition gives $\boldsymbol{v}^h(\boldsymbol{m}_E)=\boldsymbol{0}$ in the barycenters of the boundary faces. It follows that

$$\sum_{K\in\mathcal{T}^h}\int_K\nabla\cdot\boldsymbol{v}^h(\boldsymbol{x})\,d\boldsymbol{x}=0.$$

Using that the restriction of $\nabla\cdot\boldsymbol{v}^h$ to a mesh cell is a constant function, one obtains with (4.85)

$$0=\frac{|\Omega|}{|K_1|}\int_{K_1}\nabla\cdot\boldsymbol{v}^h(\boldsymbol{x})\,d\boldsymbol{x}=|\Omega|\left.\nabla\cdot\boldsymbol{v}^h\right|_{K_1}.$$

Since $|\Omega| > 0$, one finds that $\nabla \cdot \boldsymbol{v}^h\big|_{K_1} = 0$ for an arbitrarily chosen mesh cell K_1. Hence, $\boldsymbol{v}^h$ is divergence-free in the interior of each mesh cell. ∎

Remark 4.63 (Functions from V^h_{div} Are Not Weakly Divergence-free) It is $V^h_{\mathrm{div}} \not\subset H_{\mathrm{div}}(\Omega)$. Functions from $H_{\mathrm{div}}(\Omega)$ belong to $H(\mathrm{div}, \Omega)$ and it is known from Lemma 3.66 that a necessary condition for such functions is that $\boldsymbol{v}^h \cdot \boldsymbol{n}_E$ is continuous on the faces. This property is generally not given for functions from $V^h = P_1^{\mathrm{nc}}$. In particular, generally $\mathrm{div}^h \boldsymbol{v}^h \not\in L^2(\Omega)$ for $\boldsymbol{v}^h \in V^h$, see Remark 3.55 for the definition of div^h.

Thus, the divergence-free property of Lemma 4.62 is not the property of the discrete solution to be weakly divergence-free. It is a weaker property, which is reflected, e.g., in the dependency of the bound of the velocity error (4.75) on the pressure. This dependency is not given for conforming weakly divergence-free finite elements, see (4.48). A way for removing this dependency and computing weakly divergence-free solutions based on the Crouzeix–Raviart pair of finite element spaces is presented in Sect. 4.6.2. □

Remark 4.64 (Scaled Stokes Equations) Error estimate for the scaled Stokes equations (4.44) can be derived from (4.75) and (4.76) with a scaling argument as already explained in Remark 4.33. One gets

$$\left\|\boldsymbol{u} - \boldsymbol{u}^h\right\|_{V^h} \le Ch\left(|\boldsymbol{u}|_{H^2(\Omega)} + \frac{1}{\nu}\|\nabla p\|_{L^2(\Omega)}\right) \tag{4.86}$$

and

$$\left\|p - p^h\right\|_{L^2(\Omega)} \le Ch\left(1 + \frac{1}{\beta^h_{\mathrm{is}}}\right)\left(\nu\,|\boldsymbol{u}|_{H^2(\Omega)} + \|\nabla p\|_{L^2(\Omega)}\right). \tag{4.87}$$

Under the assumptions of Theorem 4.61, it holds

$$\left\|\boldsymbol{u} - \boldsymbol{u}^h\right\|_{L^2(\Omega)} \le Ch^2\left(|\boldsymbol{u}|_{H^2(\Omega)} + \frac{1}{\nu}\|\nabla p\|_{L^2(\Omega)}\right). \tag{4.88}$$

□

Example 4.65 (Scaled Stokes Equations with Analytic Solution for Inf-Sup Stable Lowest Order Pairs of Finite Element Spaces) The scaled Stokes equations (4.44) were considered with the prescribed solution given in Example D.3. Simulations were performed with the P_1^{nc}/P_0 and the Q_1^{rot}/Q_0 pairs of finite element spaces on unstructured and irregular grids, respectively, see Fig. 4.2 for the coarsest grids.

The results are presented in Fig. 4.10. The dependency of the errors on the value of the viscosity can be observed very well. In contrast to the results for the

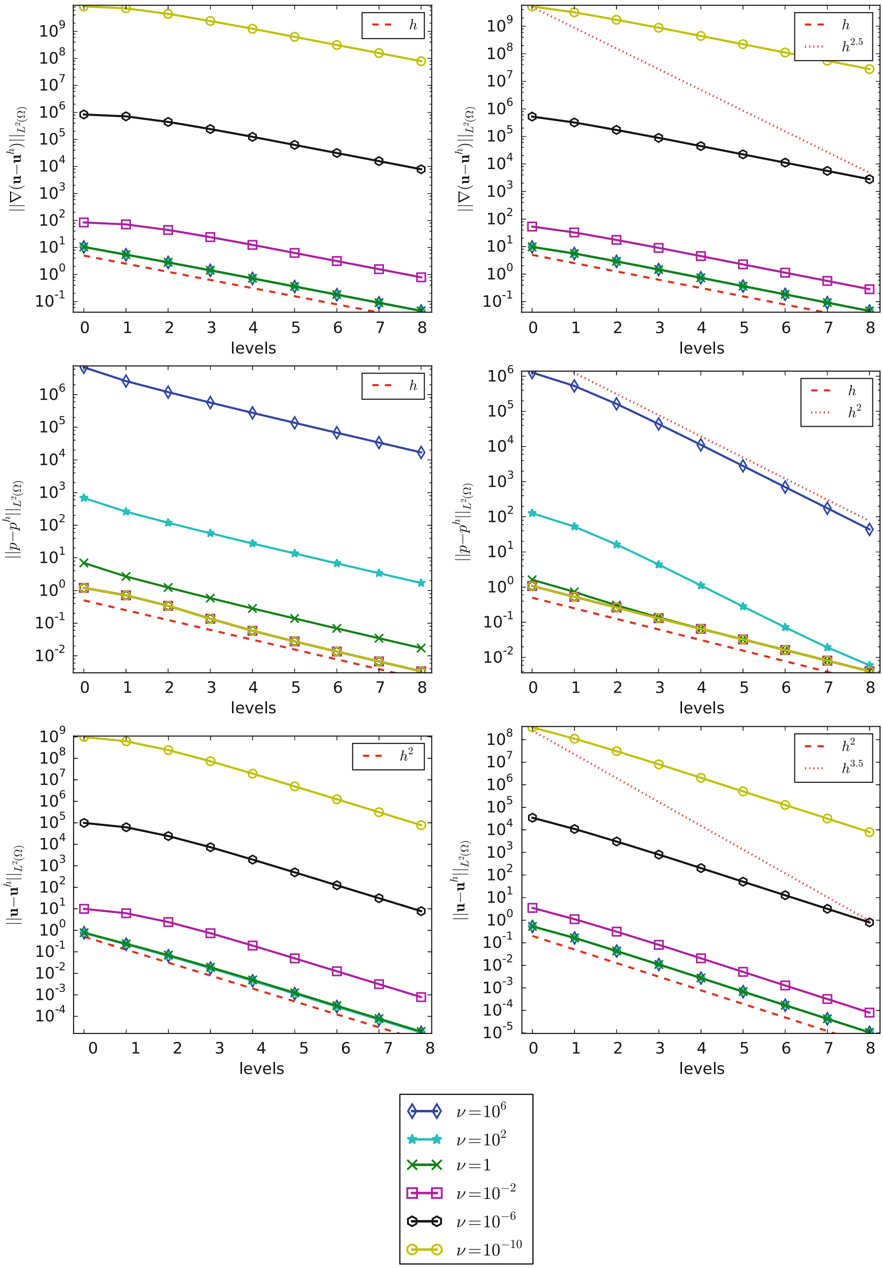

Fig. 4.10 Example 4.65. Convergence of the errors for the scaled Stokes problem and the pair of finite element spaces P_1^{nc}/P_0 (*left*) and Q_1^{rot}/Q_0 (*right*)

conforming P_2/P_1 pair of spaces, see Fig. 4.7, there is usually no higher order of error reduction for large errors. The only exception is the pressure error in $L^2(\Omega)$ for Q_1^{rot}/Q_0. This example already shows that for the pair P_1^{nc}/P_0 the dependency of the error bounds (4.86)–(4.88) on the viscosity is sharp. □

4.3 Implementation of Finite Element Methods

Remark 4.66 (Derivation of the Algebraic Saddle Point Problem) In order to derive a linear algebraic system from (4.17), the spaces V^h and Q^h are equipped with a basis. A standard approach for choosing the basis of the vector-valued velocity space is as follows

$$\begin{aligned} V^h &= \operatorname{span}\{\boldsymbol{\phi}_i^h\}_{i=1}^{3N_v} \\ &= \operatorname{span}\left\{\left\{\begin{pmatrix}\phi_i^h\\0\\0\end{pmatrix}\right\}_{i=1}^{N_v} \cup \left\{\begin{pmatrix}0\\\phi_i^h\\0\end{pmatrix}\right\}_{i=1}^{N_v} \cup \left\{\begin{pmatrix}0\\0\\\phi_i^h\end{pmatrix}\right\}_{i=1}^{N_v}\right\}, \end{aligned}$$

i.e., each basis function does not vanish in one component only. Here, N_v is the number of unknowns (degrees of freedom, d.o.f.) for one component of the velocity. The pressure space is

$$Q^h = \operatorname{span}\{\psi_i^h\}_{i=1}^{N_p},$$

where N_p is the number of pressure d.o.f.

Hence, one has the unique representation

$$\boldsymbol{u}^h = \sum_{j=1}^{3N_v} u_j^h \boldsymbol{\phi}_j^h, \quad p^h = \sum_{j=1}^{N_p} p_j^h \psi_j^h, \tag{4.89}$$

with unknown real coefficients $\underline{u} = \left(u_j^h\right)_{j=1}^{3N_v}$ and $\underline{p} = \left(p_j^h\right)_{j=1}^{N_p}$. Inserting (4.89) in (4.17) and testing with each basis function separately gives the linear system of equations

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{0} \end{pmatrix}, \tag{4.90}$$

with

$$(A)_{ij} = a_{ij} = \sum_{K\in\mathcal{T}^h} \left(\nabla\boldsymbol{\phi}_j^h, \nabla\boldsymbol{\phi}_i^h\right)_K, \ i,j = 1,\ldots,3N_v, \tag{4.91}$$

$$(B)_{ij} = b_{ij} = -\sum_{K\in\mathcal{T}^h} \left(\nabla\cdot\boldsymbol{\phi}_j^h, \psi_i^h\right)_K, \ i = 1,\ldots,N_p,\ j = 1,\ldots,3N_v, \tag{4.92}$$

$$(\underline{f})_i = f_i = \sum_{K\in\mathcal{T}^h} \left(\boldsymbol{f}, \boldsymbol{\phi}_i^h\right)_K, \ i = 1,\ldots,3N_v.$$

The expression (4.91) vanishes if the non-zero component of $\boldsymbol{\phi}_i^h$ and $\boldsymbol{\phi}_j^h$ is different, e.g., let $\boldsymbol{\phi}_j^h = (\phi_j^h, 0, 0)^T$ and $\boldsymbol{\phi}_i^h = (0, \phi_i^h, 0)^T$, then

$$\left(\nabla\boldsymbol{\phi}_j^h, \nabla\boldsymbol{\phi}_i^h\right) = \sum_{K\in\mathcal{T}^h}\int_K \begin{pmatrix} \partial_x\phi_j^h & \partial_y\phi_j^h & \partial_z\phi_j^h \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} : \begin{pmatrix} 0 & 0 & 0 \\ \partial_x\phi_i^h & \partial_y\phi_i^h & \partial_z\phi_i^h \\ 0 & 0 & 0 \end{pmatrix} d\boldsymbol{x} = 0.$$

Similarly, one finds in the case that the non-zero components of $\boldsymbol{\phi}_j^h$ and $\boldsymbol{\phi}_i^h$ are the same that one gets the same value, independent of the component. Hence, the matrix A has the block structure

$$A = \begin{pmatrix} A_{11} & 0 & 0 \\ 0 & A_{11} & 0 \\ 0 & 0 & A_{11} \end{pmatrix}. \tag{4.93}$$

□

Lemma 4.67 (Properties of the Matrix A**)** *Let* $(a^h(\cdot,\cdot))^{1/2}$ *define a norm in* V^h. *Then, the matrix* $A \in \mathbb{R}^{3N_v\times 3N_v}$ *is symmetric and positive definite.*

Proof The symmetry of A follows directly from definition (4.91).

Let $\boldsymbol{v}^h \in V^h$ with $\boldsymbol{v}^h \neq \boldsymbol{0}$ be arbitrary. Then, there is a unique representation $\boldsymbol{v}^h = \sum_{i=1}^{3N_v} v_i\boldsymbol{\phi}_i^h$ with $\underline{v} \in \mathbb{R}^{3N_v}\setminus\{\underline{0}\}$. Since $(a^h(\cdot,\cdot))^{1/2}$ defines a norm in V^h, it holds

$$\begin{aligned} 0 &< \left\|\boldsymbol{v}^h\right\|_{V^h}^2 = a^h\left(\boldsymbol{v}^h, \boldsymbol{v}^h\right) = \sum_{K\in\mathcal{T}^h}\left(\nabla\boldsymbol{v}^h, \nabla\boldsymbol{v}^h\right)_K \\ &= \sum_{K\in\mathcal{T}^h}\left(\sum_{j=1}^{3N_v} v_j\nabla\boldsymbol{\phi}_j^h, \sum_{i=1}^{3N_v} v_i\nabla\boldsymbol{\phi}_i^h\right)_K = \sum_{i,j=1}^{3N_v} v_j v_i \sum_{K\in\mathcal{T}^h}\left(\nabla\boldsymbol{\phi}_j^h, \nabla\boldsymbol{\phi}_i^h\right)_K \\ &= \underline{v}^T A\underline{v}. \end{aligned}$$

Because of the isomorphism between V^h and $\mathbb{R}^{3N_v}$, this inequality is just the positive definiteness of A. ■

Lemma 4.68 (Properties of the Saddle Point Matrices) *Let the matrices $A \in \mathbb{R}^{3N_v \times 3N_v}$ and $B \in \mathbb{R}^{N_p \times 3N_v}$ be the matrices which are assembled with a pair of finite element spaces that satisfies the discrete inf-sup condition (3.51). Let A be symmetric and positive definite, then the matrix*

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \tag{4.94}$$

is symmetric and indefinite, whereas the matrix

$$\begin{pmatrix} A & B^T \\ -B & 0 \end{pmatrix} \tag{4.95}$$

is positive semi-definite.

Proof The symmetry of (4.94) is obvious. It is

$$\left(\underline{v}, \underline{q}\right) \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{v} \\ \underline{q} \end{pmatrix} = \underline{v}^T A \underline{v} + \underline{v}^T B^T \underline{q} + \underline{q}^T B \underline{v} = \underline{v}^T A \underline{v} + 2\underline{q}^T B \underline{v},$$

since $\underline{v}^T B^T \underline{q} = \left(\underline{v}^T B^T \underline{q}\right)^T = \underline{q}^T B \underline{v}$ because it is a real number. Hence, it is for $\underline{v} \neq \underline{0}$

$$\left(\underline{v}, \underline{0}\right) \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{v} \\ \underline{0} \end{pmatrix} = \underline{v}^T A \underline{v} > 0.$$

Now, let $\underline{v} = -A^{-1} B^T \underline{q}$. The matrix B^T represents the discrete gradient operator, which is an isomorphism between Q^h and a subspace of $\left(V^h\right)'$, see Lemma 3.12. In the finite-dimensional case, the dual space of V^h can be identified with V^h, hence the image of the discrete gradient is a subspace of V^h. Thus, $B^T \underline{q} \neq \underline{0}$ if $\underline{q} \neq \underline{0}$. Since A is symmetric and positive definite, it follows that $A^{-1/2} B^T \underline{q} \neq \underline{0}$ and hence

$$\begin{aligned} \underline{v}^T A \underline{v} + 2\underline{q}^T B \underline{v} &= \underline{q}^T B A^{-1} B^T \underline{q} - 2\underline{q}^T B A^{-1} B^T \underline{q} = -\underline{q}^T B A^{-1} B^T \underline{q} \\ &= -\left(A^{-1/2} B^T \underline{q}\right)^T A^{-1/2} B^T \underline{q} = -\left\| A^{-1/2} B^T \underline{q} \right\|_2^2 < 0. \end{aligned}$$

Hence, the matrix (4.94) is indefinite.

For the matrix (4.95), one obtains

$$\left(\underline{v}, \underline{q}\right) \begin{pmatrix} A & B^T \\ -B & 0 \end{pmatrix} \begin{pmatrix} \underline{v} \\ \underline{q} \end{pmatrix} = \underline{v}^T A \underline{v} + \underline{v}^T B^T \underline{q} - \underline{q}^T B \underline{v} = \underline{v}^T A \underline{v}.$$

If $\underline{v} \neq \underline{0}$, then this bilinear form takes a positive value, otherwise is zero for all $\underline{q}$. ■

Remark 4.69 (Deformation Tensor Formulation of the Viscous Term) The equivalence of the Laplacian formulation of the viscous term and the deformation tensor formulation was derived under the condition that the velocity solution is (weakly) divergence-free, see Remark 2.20. However, finite element velocity functions are generally only discretely divergence-free, hence in the context of finite element methods, one gets in general different solutions with the different formulations. The weak formulation of with deformation tensor is given in (4.5).

Since velocity finite element basis functions are not discretely divergence-free, the use of the deformation tensor formulation leads also to a different structure of the matrix block A compared with (4.93):

$$\begin{aligned}
&\left(2\mathbb{D}\left(\boldsymbol{\phi}_j^h\right), \mathbb{D}\left(\boldsymbol{\phi}_i^h\right)\right) \\
&= 2\left(\frac{\nabla\boldsymbol{\phi}_j^h + \left(\nabla\boldsymbol{\phi}_j^h\right)^T}{2}, \frac{\nabla\boldsymbol{\phi}_i^h + \left(\nabla\boldsymbol{\phi}_i^h\right)^T}{2}\right) \\
&= \frac{1}{2}\left(\sum_{K\in\mathcal{T}^h} \left(\nabla\boldsymbol{\phi}_j^h, \nabla\boldsymbol{\phi}_i^h\right)_K + \left(\nabla\boldsymbol{\phi}_j^h, \left(\nabla\boldsymbol{\phi}_i^h\right)^T\right)_K + \left(\left(\nabla\boldsymbol{\phi}_j^h\right)^T, \nabla\boldsymbol{\phi}_i^h\right)_K\right. \\
&\quad \left. + \left(\left(\nabla\boldsymbol{\phi}_j^h\right)^T, \left(\nabla\boldsymbol{\phi}_i^h\right)^T\right)_K\right) \\
&= \sum_{K\in\mathcal{T}^h} \left(\nabla\boldsymbol{\phi}_j^h, \nabla\boldsymbol{\phi}_i^h\right)_K + \left(\left(\nabla\boldsymbol{\phi}_j^h\right)^T, \nabla\boldsymbol{\phi}_i^h\right)_K, \quad i,j = 1,\ldots,3N_v.
\end{aligned}$$

The second term on the right-hand side does not vanish any longer if the non-zero components of $\boldsymbol{\phi}_i^h$ and $\boldsymbol{\phi}_j^h$ are different. Consider again $\boldsymbol{\phi}_j^h = (\phi_j^h, 0, 0)^T$ and $\boldsymbol{\phi}_i^h = (0, \phi_i^h, 0)^T$, then

$$\begin{aligned}
\left(\left(\nabla\boldsymbol{\phi}_j^h\right)^T, \nabla\boldsymbol{\phi}_i^h\right) &= \sum_{K\in\mathcal{T}^h} \int_K \begin{pmatrix} \partial_x\phi_j^h & 0 & 0 \\ \partial_y\phi_j^h & 0 & 0 \\ \partial_z\phi_j^h & 0 & 0 \end{pmatrix} : \begin{pmatrix} 0 & 0 & 0 \\ \partial_x\phi_i^h & \partial_y\phi_i^h & \partial_z\phi_i^h \\ 0 & 0 & 0 \end{pmatrix} d\boldsymbol{x} \\
&= \sum_{K\in\mathcal{T}^h} \int_K \partial_y\phi_j^h \partial_x\phi_i^h(\boldsymbol{x})\, d\boldsymbol{x}.
\end{aligned}$$

Altogether, one gets a symmetric matrix without zero blocks. The matrix for the deformation tensor formulation has the form

$$A = \begin{pmatrix} A_{11} & 0 & 0 \\ 0 & A_{11} & 0 \\ 0 & 0 & A_{11} \end{pmatrix} + \begin{pmatrix} \tilde{A}_{11} & A_{12} & A_{13} \\ A_{12}^T & \tilde{A}_{22} & A_{23} \\ A_{13}^T & A_{23}^T & \tilde{A}_{33} \end{pmatrix} = \begin{pmatrix} A_{11} + \tilde{A}_{11} & A_{12} & A_{13} \\ A_{12}^T & A_{11} + \tilde{A}_{22} & A_{23} \\ A_{13}^T & A_{23}^T & A_{11} + \tilde{A}_{33} \end{pmatrix}.$$

Consequently, using the deformation tensor formulation requires to store considerably more matrix entries than for the Laplacian or gradient formulation. In addition, matrix-vector products are computationally more expensive for the deformation tensor formulation.

For Stokes flows, stationary, and laminar Navier–Stokes flows, one finds in the literature most often the use of the Laplacian or gradient formulation of the viscous term. □

Remark 4.70 (The Finite Element Pressure Space Q^h and Standard Finite Element Space Q^h_{std}) A standard implementation of finite element methods for incompressible flows uses for the pressure space a standard finite element space as presented in Appendix B.2 with the standard nodal basis. However, this basis is generally not in $L^2_0(\Omega)$, e.g., the integral on Ω of the standard basis function of P_1, the hat function, is positive. Thus, one cannot expect that a linear combination of such basis functions belongs generally to $L^2_0(\Omega)$.

Continuity Equation It will be discussed first that the standard basis functions can be used as test functions in the finite element continuity equation although they are not from $L^2_0(\Omega)$. The finite element space spanned by the standard basis is denoted by Q^h_{std}, it will be assumed that $1 \in Q^h_{\mathrm{std}}$, and it is $Q^h = Q^h_{\mathrm{std}} \cap L^2_0(\Omega)$. Let ψ^h be such a standard basis function. Then there is a $C_\psi \in \mathbb{R}$,

$$C_\psi = -\frac{\int_\Omega \psi^h(\boldsymbol{x})\, d\boldsymbol{x}}{\int_\Omega d\boldsymbol{x}} = -\frac{\int_\Omega \psi^h(\boldsymbol{x})\, d\boldsymbol{x}}{|\Omega|}, \tag{4.96}$$

such that $\psi^h + C_\psi \in L^2_0(\Omega)$, i.e., $\psi^h + C_\psi \in Q^h$.

Consider first conforming velocity finite element spaces $V^h \subset V$. It follows from (4.17) with integration by parts that

$$\begin{aligned} 0 &= b(\boldsymbol{u}^h, \psi^h + C_\psi) = -\left(\nabla\cdot\boldsymbol{u}^h, \psi^h + C_\psi\right) = -\left(\nabla\cdot\boldsymbol{u}^h, \psi^h\right) - \left(\nabla\cdot\boldsymbol{u}^h, C_\psi\right) \\ &= -\left(\nabla\cdot\boldsymbol{u}^h, \psi^h\right) - \int_{\partial\Omega} C_\psi \boldsymbol{u}^h\cdot\boldsymbol{n}\, ds + \left(\boldsymbol{u}^h, \nabla C_\psi\right) = -\left(\nabla\cdot\boldsymbol{u}^h, \psi^h\right). \end{aligned}$$

The boundary integral vanishes because of the homogeneous Dirichlet boundary condition fulfilled by $\boldsymbol{u}^h$. This equality reveals that testing with basis functions from Q^h_{std} imposes the property that $\boldsymbol{u}^h$ is discretely divergence-free. Thus, the basis functions from Q^h_{std} can be used as test functions.

For non-conforming finite element spaces, one obtains with integration by parts

$$\begin{aligned} 0 = b^h(\boldsymbol{u}^h, \psi^h + C_\psi) &= - \sum_{K\in\mathcal{T}^h} \left(\nabla\cdot\boldsymbol{u}^h, \psi^h + C_\psi\right)_K \\ &= - \sum_{K\in\mathcal{T}^h} \left[\left(\nabla\cdot\boldsymbol{u}^h, \psi^h\right)_K + \int_{\partial K} C_\psi \boldsymbol{u}^h\cdot\boldsymbol{n}_{\partial K}\, d\boldsymbol{s}\right]. \end{aligned}$$

The standard basis functions can be uses as test functions if the integrals on ∂K vanish. Thus, it must hold

$$\int_E \left[\left|\boldsymbol{u}^h\cdot\boldsymbol{n}_E\right|\right]_E(\boldsymbol{s})\, d\boldsymbol{s} = 0 \quad \forall\, E \in \mathcal{E}^h \tag{4.97}$$

for all interior faces of the triangulation and

$$\int_E \boldsymbol{u}^h\cdot\boldsymbol{n}_E(\boldsymbol{s})\, d\boldsymbol{s} = 0 \quad \forall\, E \in \overline{\mathcal{E}}^h \setminus \mathcal{E}^h \tag{4.98}$$

for all faces on the boundary of the domain, where the jumps are defined in (3.72). These conditions are satisfied, e.g., for the Crouzeix–Raviart finite element P_1^{nc} and the mean-value-oriented Rannacher–Turek element Q_1^{rot}.

Momentum Equation Similarly, the ansatz $p^h = \sum_{j=1}^{N_{p,\mathrm{std}}} p_j\psi_j^h$ with the standard basis functions can be used in the momentum equation. Again, there is a constant such that $p^h + C_p \in Q^h$. For the momentum equation, one gets with integration by parts and with the homogeneous Dirichlet conditions of the velocity test functions

$$b^h(\boldsymbol{v}^h, p^h + C_p) = - \sum_{K\in\mathcal{T}^h} \left(\nabla\cdot\boldsymbol{v}^h, p^h + C_p\right)_K = - \sum_{K\in\mathcal{T}^h} \left(\nabla\cdot\boldsymbol{v}^h, p^h\right)_K,$$

where for non-conforming finite element spaces one needs the assumptions (4.97) and (4.98). Examples of non-conforming spaces satisfying these conditions were already given in the discussion of the continuity equation.

In summary, it is possible to use not normalized ansatz and test functions for the discrete pressure in the evaluation of the bilinear form $b^h(\cdot,\cdot)$.

The use of a standard nodal basis for the finite element pressure has the following advantages:

- The application of a pressure finite element space which is in $L_0^2(\Omega)$ requires the use of a non-standard basis. This basis will be less local than the standard basis. Consequently, the matrices become denser and the numerical cost of matrix operations increases.

- A general code for solving incompressible flow problems should be able to handle also others than Dirichlet boundary conditions. All boundary conditions which include conditions on the stress tensor on the boundary, and therefore conditions on the pressure on the boundary, do not need an additional normalization of the pressure. Boundary conditions of such type are the do-nothing condition (2.37) or slip boundary conditions like (2.34). In these cases, it is just required that $Q^h \subset L^2(\Omega)$ and the application of a pressure finite element space with a standard basis is natural.

□

Lemma 4.71 (Dimension of Q^h and Q^h_{std}) *Let* $Q^h_{\mathrm{std}} \subset L^2(\Omega)$ *be a standard finite element space with* $1 \in Q^h_{\mathrm{std}}$ *and let* $Q^h = Q^h_{\mathrm{std}} \cap L^2_0(\Omega)$. *Then it holds* $\dim\left(Q^h_{\mathrm{std}}\right) = \dim\left(Q^h\right) + 1 = N_p + 1$.

Proof Let $\{\psi_i^h\}_{i=1}^N$ be a basis of Q^h_{std}. Since $1 \in Q^h_{\mathrm{std}}$, there is a unique representation

$$1 = \sum_{i=1}^N \beta_i \psi_i^h,$$

where at least one β_i is not equal to zero. Without less of generality, it can be assumed that $\beta_1 \neq 0$. Then, a new basis of Q^h_{std} is given by

$$\{1\} \cup \{\psi_i^h\}_{i=2}^N. \tag{4.99}$$

i) $\dim\left(Q^h\right)$ *is at most* $\dim\left(Q^h_{\mathrm{std}}\right) - 1$. By construction, it holds that $Q^h \subset Q^h_{\mathrm{std}}$ and consequently, it follows that $\dim\left(Q^h\right) \le \dim\left(Q^h_{\mathrm{std}}\right)$. Since both spaces are finite-dimensional, they are isometric isomorph to $\mathbb{R}^n$ for some n, see Definition A.3. If the dimensions of Q^h and Q^h_{std} would be the same, then they would be both isometric isomorph to $\mathbb{R}^N$ and consequently the spaces would be isometric isomorph to each other. That means, each element of one space can be identified with an element from the other space. Since all functions of Q^h belong to Q^h_{std}, and thus can be identified with themself, and $1 \in Q^h_{\mathrm{std}}$ but $1 \notin Q^h$, the spaces cannot be isometric isomorph. Therefore, it follows that $\dim\left(Q^h\right) < \dim\left(Q^h_{\mathrm{std}}\right)$.

ii) $\dim\left(Q^h\right)$ *is at least* $\dim\left(Q^h_{\mathrm{std}}\right) - 1$. For each ψ_i^h there is a constant C_i such that the integral mean value of $\psi_i^h + C_i$ vanishes, see (4.96), and hence $\psi_i^h + C_i \in L^2_0(\Omega)$, $i = 2, \ldots, N$. Since $1 \in Q^h_{\mathrm{std}}$, it follows that $C_i \in Q^h_{\mathrm{std}}$ and consequently that $\psi_i^h + C_i \in Q^h_{\mathrm{std}}$. Because the integral mean value of $\psi_i^h + C_i$ is zero, it holds even $\psi_i^h + C_i \in Q^h$. It will be shown with a proof by contradiction that the functions $\{\psi_i^h + C_i\}$, $i = 2, \ldots, N$, are linearly independent. Without loss of generality, assume that $\psi_2^h + C_2$ can be represented by a linear combination of the functions $\{\psi_i^h + C_i\}$, $i = 3, \ldots, N$:

$$\psi_2^h + C_2 = \sum_{i=3}^{N} \gamma_i \left(\psi_i^h + C_i\right) \quad \Longleftrightarrow \quad \psi_2^h = \sum_{i=3}^{N} \gamma_i \psi_i^h + C \cdot 1,$$

where all constants $\gamma_i C_i$ were absorbed by the constant $C \in \mathbb{R}$. Thus, there is a representation of ψ_2^h by a linear combination of $\{\psi_i^h + C_i\}$, $i = 3, \ldots, N$, and 1. This result is a contradiction to the fact that (4.99) is a basis of Q_{std}^h, i.e., these elements are in particular linearly independent. In this way $\left(\dim\left(Q_{\mathrm{std}}^h\right) - 1\right)$ linearly independent elements were found in Q^h, which proves that $\dim\left(Q^h\right) \geq \dim\left(Q_{\mathrm{std}}^h\right) - 1$. ■

Remark 4.72 (Numbers of Matrix Entries per Row or Column) In the case of conforming finite element discretizations, the number of matrix entries per row or column depends usually on the local grid, e.g., on the number of mesh cells where a vertex belongs to.

The situation is different for lowest order non-conforming finite element spaces since in these spaces, the degrees of freedom for the velocity are defined only on the faces of the mesh cells and the degrees of freedom for the pressure in the interior of the mesh cells. Considering an arbitrary face and the two mesh cells sharing this face. Then, the degree of freedom defined on the considered face leads to non-zero entries only in combination with the degrees of freedom defined on the faces and the interior of the two mesh cells. One finds that in each row/column of the matrix A_{11} from (4.93) there are at most $5/7$ entries for P_1^{nc} in two/three dimensions and $7/11$ entries for Q_1^{rot} in two/three dimensions. Similarly, the constant pressure degrees of freedom lead to at most $6/12$ non-zero entries in each row of the matrix B for P_1^{nc}/P_0 and $8/18$ non-zero entries for Q_1^{rot}/Q_0 in two and three dimensions, respectively. □

4.4 Residual-Based A Posteriori Error Analysis

Remark 4.73 (Goals of A Posteriori Error Estimators) So far, so called a priori error estimates were proved, e.g., in Corollary 4.30. The right-hand side of such estimates depends on the unknown solution of the continuous problem and on an unknown constant.

The goals of a posteriori error estimates are twofold. First, they should provide computable estimates of errors between a computed solution $\left(\boldsymbol{u}^h, p^h\right)$ and the unknown solution $(\boldsymbol{u}, p)$ of the continuous problem (4.2). Usually, the errors are measured in norms of Sobolev spaces defined on Ω. That means, estimates of global errors from above have to be derived, which have the form, e.g., for the velocity

$$\left\|\boldsymbol{u} - \boldsymbol{u}^h\right\|_{\Omega} \leq C\eta. \tag{4.100}$$

In (4.100), η is a quantity which is computable with the information available in the numerical solution process and C is a positive constant which should be independent of the mesh width and the solution and for which one should have an idea of its order of magnitude.

The second task of a posteriori error estimates consists in controlling an adaptive mesh refinement. The rationale behind this idea is that the error of the computed solution to the solution of the continuous problem can be reduced best, or at least significantly, if the mesh in those subregions of the domain is refined, where the local error is largest. Then, a local a posteriori error estimate should identify these subregions. To this end, a local lower estimate from below of the form

$$\eta_K \le C \left\| \boldsymbol{u} - \boldsymbol{u}^h \right\|_{\omega(K)} \tag{4.101}$$

is necessary, where $\omega(K)$ denotes a small neighborhood of a mesh cell K and η_K is a computable quantity. For (4.101), it has to be proved that the positive constant C can be bounded from below and above independently of K. Estimate (4.101) tells that in subregions where the local error estimate η_K is large, also the local error is large.

There are different ways for computing a posteriori error estimates, see, e.g., the monographs Ainsworth and Oden (2000), Verfürth (2013). Among the most popular ones are residual-based estimates, which where presented for the Stokes equations the first time in Verfürth (1989), and estimates which are based on the solution of local problems, that were also introduced in Verfürth (1989).

A review of residual-based estimators can be found in Verfürth (2013). □

Remark 4.74 (Basic Assumptions) This section considers conforming inf-sup stable pairs of finite element spaces, which are defined on a quasi-uniform family of triangulations. □

Lemma 4.75 (Estimate of the Supremum of the Bilinear Form) *For all* $(\boldsymbol{w}, r) \in V \times Q$ *it is*

$$\begin{aligned} \|\nabla \boldsymbol{w}\|_{L^2(\Omega)} + \|r\|_{L^2(\Omega)} &\le \frac{1}{\beta_{\text{is,Bab}}} \sup_{\substack{(\boldsymbol{v},q)\in V\times Q \\ (\boldsymbol{v},q)\neq(\boldsymbol{0},0)}} \frac{(\nabla \boldsymbol{w}, \nabla \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, r) + (\nabla \cdot \boldsymbol{w}, q)}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)} + \|q\|_{L^2(\Omega)}} \\ &\le \frac{1}{\beta_{\text{is,Bab}}} \left(\|\nabla \boldsymbol{w}\|_{L^2(\Omega)} + \|r\|_{L^2(\Omega)} \right). \end{aligned} \tag{4.102}$$

Proof The first estimate follows directly from the Babuška inf-sup condition (4.11). For the second estimate, one applies the Cauchy–Schwarz inequality (A.10) and (3.41), which gives

$$
\begin{aligned}
&\sup_{\substack{(\boldsymbol{v},q)\in V\times Q\\(\boldsymbol{v},q)\neq(\boldsymbol{0},0)}} \frac{(\nabla \boldsymbol{w},\nabla \boldsymbol{v})-(\nabla\cdot\boldsymbol{v},r)+(\nabla\cdot\boldsymbol{w},q)}{\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}+\|q\|_{L^2(\Omega)}}\\
&\le \sup_{\substack{(\boldsymbol{v},q)\in V\times Q\\(\boldsymbol{v},q)\neq(\boldsymbol{0},0)}} \frac{\|\nabla\boldsymbol{w}\|_{L^2(\Omega)}\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}+\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}\|r\|_{L^2(\Omega)}+\|\nabla\boldsymbol{w}\|_{L^2(\Omega)}\|q\|_{L^2(\Omega)}}{\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}+\|q\|_{L^2(\Omega)}}\\
&\le \sup_{\substack{(\boldsymbol{v},q)\in V\times Q\\(\boldsymbol{v},q)\neq(\boldsymbol{0},0)}} \frac{\left(\|\nabla\boldsymbol{w}\|_{L^2(\Omega)}+\|r\|_{L^2(\Omega)}\right)\left(\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}+\|q\|_{L^2(\Omega)}\right)}{\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}+\|q\|_{L^2(\Omega)}}\\
&= \|\nabla\boldsymbol{w}\|_{L^2(\Omega)}+\|r\|_{L^2(\Omega)}\,.
\end{aligned}
$$

■

Remark 4.76 (To Estimate (4.102)*)* Estimate (4.102) states the fact that the operator in the numerator defines an isomorphism from $V\times Q$ to $V'\times Q'$. The left estimate gives injectivity of the operator since if two different arguments would give the same results, the term in the middle of (4.102) vanishes for the difference and the left-hand side does not, which is a contradiction. The right estimate holds for all arguments and thus it just gives the boundedness of the operator. Altogether, an estimate of type (4.102) is not special for the Stokes equations but an estimate of this type holds always for well-posed linear problems, compare (Verfürth 2013, Proposition 4.1). □

Corollary 4.77 (Error Estimate with the Residual) *Let $V^h\subset V$ and $Q^h\subset Q$, let $(\boldsymbol{u},p)$ be the solution of* (4.2) *and $\left(\boldsymbol{u}^h,p^h\right)\in V^h\times Q^h$ be arbitrary, then it holds*

$$
\begin{aligned}
&\sup_{\substack{(\boldsymbol{v},q)\in V\times Q\\(\boldsymbol{v},q)\neq(\boldsymbol{0},0)}} \frac{(\boldsymbol{f},\boldsymbol{v})-(\nabla\boldsymbol{u}^h,\nabla\boldsymbol{v})+(\nabla\cdot\boldsymbol{v},p^h)-(\nabla\cdot\boldsymbol{u}^h,q)}{\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}+\|q\|_{L^2(\Omega)}}\\
&\le \left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)}+\left\|p-p^h\right\|_{L^2(\Omega)} \qquad (4.103)\\
&\le \frac{1}{\beta_{\mathrm{is,Bab}}}\sup_{\substack{(\boldsymbol{v},q)\in V\times Q\\(\boldsymbol{v},q)\neq(\boldsymbol{0},0)}} \frac{(\boldsymbol{f},\boldsymbol{v})-(\nabla\boldsymbol{u}^h,\nabla\boldsymbol{v})+(\nabla\cdot\boldsymbol{v},p^h)-(\nabla\cdot\boldsymbol{u}^h,q)}{\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}+\|q\|_{L^2(\Omega)}}.
\end{aligned}
$$

Proof Setting $\boldsymbol{w}=\boldsymbol{u}-\boldsymbol{u}^h$ and $r=p-p^h$ in (4.102) gives for the numerator

$$
\begin{aligned}
&(\nabla(\boldsymbol{u}-\boldsymbol{u}^h),\nabla\boldsymbol{v})-(\nabla\cdot\boldsymbol{v},p-p^h)+(\nabla\cdot(\boldsymbol{u}-\boldsymbol{u}^h),q)\\
&=(\boldsymbol{f},\boldsymbol{v})-(\nabla\boldsymbol{u}^h,\nabla\boldsymbol{v})+(\nabla\cdot\boldsymbol{v},p^h)-(\nabla\cdot\boldsymbol{u}^h,q). \qquad (4.104)
\end{aligned}
$$

Now, the statement of the corollary follows from (4.102). ■

Remark 4.78 (To Estimate (4.103)) The supremum in estimate (4.103) is just the definition of the norm of the residual in $V' \times Q'$. Hence estimate (4.103) states that the error $\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)} + \left\|p-p^h\right\|_{L^2(\Omega)}$ for an arbitrary pair $\left(\boldsymbol{u}^h, p^h\right) \in V^h \times Q^h$ is bounded from below and from above by the norm of the residual in the dual space. However, in practice one cannot compute this norm since the supremum is taken in an infinite-dimensional space. The goal consists now in estimating this norm with computable expressions. □

Theorem 4.79 (Global Upper, Residual-Based, a Posteriori Error Estimate for Conforming Inf-Sup Stable Finite Element Spaces) *Let $\boldsymbol{f} \in L^2(\Omega)$, let $P^h\boldsymbol{f}$ a polynomial approximation of $\boldsymbol{f}$ (which can be integrated exactly), and consider conforming finite element spaces V^h/Q^h which satisfy the discrete inf-sup condition (3.51) on a quasi-uniform family of triangulations $\{\mathcal{T}^h\}_{h>0}$. Let $(\boldsymbol{u}, p)$ be the solution of (4.2) and $\left(\boldsymbol{u}^h, p^h\right)$ be the solution of (4.17). Defining the mesh cell residual*

$$\boldsymbol{r}_K^h\left(\boldsymbol{u}^h, p^h\right) = \left(P^h\boldsymbol{f} + \Delta\boldsymbol{u}^h - \nabla p^h\right)\big|_K \quad \forall\, K \in \mathcal{T}^h, \tag{4.105}$$

the face residual

$$\boldsymbol{r}_E^h\left(\boldsymbol{u}^h, p^h\right) = \left[\left|\left(-\nabla\boldsymbol{u}^h + p^h\mathbb{I}\right)\boldsymbol{n}_E\right|\right]_E \quad \forall\, E \in \mathcal{E}^h, \tag{4.106}$$

and the local error estimator

$$\eta_K = \left(h_K^2\left\|\boldsymbol{r}_K^h\left(\boldsymbol{u}^h, p^h\right)\right\|_{L^2(K)}^2 + \left\|\nabla\cdot\boldsymbol{u}^h\right\|_{L^2(K)}^2 \right. \\ \left. + \frac{1}{2}\sum_{E\subset\partial K, E\in\mathcal{E}^h} h_E\left\|\boldsymbol{r}_E^h\left(\boldsymbol{u}^h, p^h\right)\right\|_{L^2(E)}^2\right)^{1/2}, \tag{4.107}$$

then it holds the a posteriori estimate

$$\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)} + \left\|p-p^h\right\|_{L^2(\Omega)} \\ \le C\left(\sum_{K\in\mathcal{T}^h}\eta_K^2 + h_K^2\left\|\boldsymbol{f}-P^h\boldsymbol{f}\right\|_{L^2(K)}^2\right)^{1/2}, \tag{4.108}$$

where the constant does not depend on the solution and on the mesh width.

Proof Subtracting the finite element equation (4.17) from the weak form of the Stokes equations (4.2) and using $\nabla\cdot\boldsymbol{u} = 0$ gives the error equation

$$\left(\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right), \nabla\boldsymbol{v}^h\right) - \left(\nabla\cdot\boldsymbol{v}^h, p-p^h\right) + \left(\nabla\cdot\left(\boldsymbol{u}-\boldsymbol{u}^h\right), q^h\right) = 0$$

for all $\left(\boldsymbol{v}^h, \boldsymbol{q}^h\right) \in V^h \times Q^h$. Using (4.104) and this equation, one gets by applying integration by parts and using $\nabla \cdot \boldsymbol{u} = 0$

$$
\begin{aligned}
&(\boldsymbol{f}, \boldsymbol{v}) - (\nabla \boldsymbol{u}^h, \nabla \boldsymbol{v}) + (\nabla \cdot \boldsymbol{v}, p^h) - (\nabla \cdot \boldsymbol{u}^h, q) \\
&= \left(\nabla\left(\boldsymbol{u} - \boldsymbol{u}^h\right), \nabla\left(\boldsymbol{v} - \boldsymbol{v}^h\right)\right) - \left(\nabla \cdot \left(\boldsymbol{v} - \boldsymbol{v}^h\right), p - p^h\right) + \left(\nabla \cdot \left(\boldsymbol{u} - \boldsymbol{u}^h\right), q - q^h\right) \\
&= \left(\nabla \boldsymbol{u}, \nabla\left(\boldsymbol{v} - \boldsymbol{v}^h\right)\right) - \left(\nabla \cdot \left(\boldsymbol{v} - \boldsymbol{v}^h\right), p\right) \\
&\quad - \sum_{K \in \mathcal{T}^h} \int_{\partial K} \left(\nabla \boldsymbol{u}^h \boldsymbol{n}_{\partial K}\right) \cdot \left(\boldsymbol{v} - \boldsymbol{v}^h\right)\, ds + \sum_{K \in \mathcal{T}^h} \left(\Delta \boldsymbol{u}^h, \boldsymbol{v} - \boldsymbol{v}^h\right)_K \\
&\quad + \sum_{K \in \mathcal{T}^h} \int_{\partial K} p^h \left(\boldsymbol{v} - \boldsymbol{v}^h\right) \cdot \boldsymbol{n}_{\partial K}\, ds - \sum_{K \in \mathcal{T}^h} \left(\nabla p^h, \boldsymbol{v} - \boldsymbol{v}^h\right)_K - \left(\nabla \cdot \boldsymbol{u}^h, q - q^h\right)
\end{aligned}
$$

for all $(\boldsymbol{v}, q) \in V \times Q$ and all $\left(\boldsymbol{v}^h, \boldsymbol{q}^h\right) \in V^h \times Q^h$. Taking now $\boldsymbol{v}^h = P^h_{\mathrm{Cle}} \boldsymbol{v}$ the Clément interpolant of $\boldsymbol{v}$ which preserves homogeneous Dirichlet boundary conditions, see Remark C.22, using that $(\boldsymbol{u}, p)$ solves the Stokes equations, writing the integrals on the faces with jumps, and using that $\boldsymbol{u}^h$ is discretely divergence-free leads to

$$
\begin{aligned}
&(\boldsymbol{f}, \boldsymbol{v}) - (\nabla \boldsymbol{u}^h, \nabla \boldsymbol{v}) + (\nabla \cdot \boldsymbol{v}, p^h) - (\nabla \cdot \boldsymbol{u}^h, q) \\
&= \sum_{K \in \mathcal{T}^h} \left[\left(P^h \boldsymbol{f} + \Delta \boldsymbol{u}^h - \nabla p^h, \boldsymbol{v} - P^h_{\mathrm{Cle}} \boldsymbol{v}\right)_K + \left(\boldsymbol{f} - P^h \boldsymbol{f}, \boldsymbol{v} - P^h_{\mathrm{Cle}} \boldsymbol{v}\right)_K - \left(\nabla \cdot \boldsymbol{u}^h, q\right)_K \right] \\
&\quad + \sum_{E \in \mathcal{E}^h} \left(\left[\left|\left(-\nabla \boldsymbol{u}^h + p^h \mathbb{I}\right) \boldsymbol{n}_E \right|\right]_E, \boldsymbol{v} - P^h_{\mathrm{Cle}} \boldsymbol{v}\right)_E,
\end{aligned}
\tag{4.109}
$$

where the jumps are defined in Remark 3.64. Note that for edges on the Dirichlet boundary it is $\boldsymbol{v} = P^h_{\mathrm{Cle}} \boldsymbol{v}$. In the next step, all terms are estimated with the Cauchy–Schwarz inequality (A.10) and the interpolation estimate (C.7)

$$
\begin{aligned}
&\sum_{K \in \mathcal{T}^h} \left(P^h \boldsymbol{f} + \Delta \boldsymbol{u}^h - \nabla p^h, \boldsymbol{v} - P^h_{\mathrm{Cle}} \boldsymbol{v}\right)_K \\
&\qquad \le C \sum_{K \in \mathcal{T}^h} h_K \left\| P^h \boldsymbol{f} + \Delta \boldsymbol{u}^h - \nabla p^h \right\|_{L^2(K)} \left\| \nabla \boldsymbol{v} \right\|_{L^2(K)} \\
&\qquad \le C \left(\sum_{K \in \mathcal{T}^h} h_K^2 \left\| \boldsymbol{r}_K^h \left(\boldsymbol{u}^h, p^h\right) \right\|_{L^2(K)}^2 \right)^{1/2} \left\| \nabla \boldsymbol{v} \right\|_{L^2(\Omega)},
\end{aligned}
$$

$$
\sum_{K \in \mathcal{T}^h} \left(\boldsymbol{f} - P^h \boldsymbol{f}, \boldsymbol{v} - P^h_{\mathrm{Cle}} \boldsymbol{v}\right)_K \le C \left(\sum_{K \in \mathcal{T}^h} h_K^2 \left\| \boldsymbol{f} - P^h \boldsymbol{f} \right\|_{L^2(K)}^2 \right)^{1/2} \left\| \nabla \boldsymbol{v} \right\|_{L^2(\Omega)},
$$

$$
\sum_{K \in \mathcal{T}^h} \left(\nabla \cdot \boldsymbol{u}^h, q\right)_K \le \left(\sum_{K \in \mathcal{T}^h} \left\| \nabla \cdot \boldsymbol{u}^h \right\|_{L^2(K)}^2 \right)^{1/2} \left\| q \right\|_{L^2(\Omega)},
$$

and with the interpolation estimate (C.24)

$$\sum_{E\in\mathcal{E}^h} \left(\left[\left|\left(-\nabla \boldsymbol{u}^h + p^h \mathbb{I}\right) \boldsymbol{n}_E\right|\right]_E, \boldsymbol{v} - P_{\mathrm{Cle}}^h \boldsymbol{v}\right)_E$$

$$\leq C \left(\sum_{E\in\mathcal{E}^h} h_E \left\| \boldsymbol{r}_E^h \left(\boldsymbol{u}^h, p^h\right) \right\|_{L^2(E)}^2 \right)^{1/2} \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}.$$

Inserting all estimates in (4.109) and observing that interior faces belong to two mesh cells leads to

$$\begin{aligned}
&(\boldsymbol{f}, \boldsymbol{v}) - (\nabla \boldsymbol{u}^h, \nabla \boldsymbol{v}) + (\nabla \cdot \boldsymbol{v}, p^h) - (\nabla \cdot \boldsymbol{u}^h, q) \\
&\leq C \Bigg(\sum_{K\in\mathcal{T}^h} h_K^2 \left\| \boldsymbol{r}_K^h \left(\boldsymbol{u}^h, p^h\right) \right\|_{L^2(K)}^2 + \left\| \nabla \cdot \boldsymbol{u}^h \right\|_{L^2(K)}^2 + h_K^2 \left\| \boldsymbol{f} - P^h \boldsymbol{f} \right\|_{L^2(K)}^2 \\
&\quad + \frac{1}{2} \sum_{E\subset\partial K, E\in\mathcal{E}^h} h_E \left\| \boldsymbol{r}_E^h \left(\boldsymbol{u}^h, p^h\right) \right\|_{L^2(E)}^2 \Bigg)^{1/2} \left(\|\nabla \boldsymbol{v}\|_{L^2(\Omega)} + \|q\|_{L^2(\Omega)} \right).
\end{aligned}$$

Now, the a posteriori error estimate (4.108) follows directly from (4.103). ■

Remark 4.80 (On the Global Upper Estimate)

- The constant on the right-hand side of (4.108) depends on the constants of the local interpolation estimates (C.7), (C.24), and on $\beta_{\mathrm{is,Bab}}^{-1}$. The inf-sup constant is related to the stability of the problem.
- For problems with non-homogeneous Dirichlet boundary conditions, instead of the Clément operator, an interpolation operator is used in the analysis that preserves these condition, usually the Scott–Zhang operator (Scott and Zhang 1990).
- In the case of do-nothing or homogeneous natural boundary conditions (2.37) on some part Γ_{donot} of the boundary, the error estimator η_K has to be extended by the term

$$\sum_{E\subset\partial K, E\subset\Gamma_{\mathrm{donot}}} h_E \left\| \boldsymbol{r}_E^h \left(\boldsymbol{u}^h, p^h\right) \right\|_{L^2(E)}^2$$

  within the parentheses, where the jump in $\boldsymbol{r}_E^h \left(\boldsymbol{u}^h, p^h\right)$ is just the difference of the boundary condition satisfied by the finite element approximation and the homogeneous boundary condition prescribed for the continuous problem.
- One can find the definition of the local estimator η_K also without the factor $1/2$ in front of the edge residuals, e.g., in Verfürth (2013, p. 243). This change of η_K changes only the constant in the estimate (4.108).

□

Remark 4.81 (Principal Way for Obtaining a Local Lower Estimate of Form (4.101)) For obtaining a local lower estimate of form (4.101), appropriate test functions are used in the continuous Stokes equations (4.2). These functions are defined with the help of mesh cell bubble functions and edge bubble functions. □

Lemma 4.82 (Local Estimates for Bubble Functions) *Let $\phi_K^h(\boldsymbol{x})$ be a cell bubble function, i.e., $\phi_K(\boldsymbol{x})$ is polynomial which is positive in the interior of K, which vanishes on ∂K, and whose support is K. Then, for all polynomials $v^h(\boldsymbol{x})$ on K the following estimates hold*

$$C^{-1} \left\| v^h \right\|_{L^2(K)}^2 \le \left(v^h, v^h \phi_K^h\right)_K \le C \left\| v^h \right\|_{L^2(K)}^2, \tag{4.110}$$

$$\begin{aligned} C^{-1} \left\| v^h \right\|_{L^2(K)} &\le \left\| v^h \phi_K^h \right\|_{L^2(K)} + h_K \left\| \nabla (v^h \phi_K^h) \right\|_{L^2(K)} \\ &\le C \left\| v^h \right\|_{L^2(K)}, \end{aligned} \tag{4.111}$$

where C is independent of v^h and h_K.

Let $\phi_E^h(\boldsymbol{x})$ be a face bubble function, i.e., $\phi_E^h(\boldsymbol{x})$ is continuous, it is a polynomial on both mesh cells K_1, K_2 which share the face E, it is positive in $\omega_E = K_1 \cup K_2$, it vanishes on the boundary of ω_E, and its support is ω_E. Then, for all polynomials v^h on K_i, $i \in \{1, 2\}$, one has the estimates

$$C^{-1} \left\| v^h \right\|_{L^2(E)}^2 \le \left(v^h, v^h \phi_E^h\right)_E \le C \left\| v^h \right\|_{L^2(E)}^2, \tag{4.112}$$

$$h_{K_i}^{-1/2} \left\| v^h \phi_E^h \right\|_{L^2(K_i)} + h_{K_i}^{1/2} \left\| \nabla (v^h \phi_E^h) \right\|_{L^2(K_i)} \le C \left\| v^h \right\|_{L^2(E)}, \tag{4.113}$$

where C is independent of v^h and h_{K_i}.

Proof The proof of these estimates can be found in Verfürth (1994), Ainsworth and Oden (2000, Theorems 2.2, 2.4), or Verfürth (2013, Proposition 1.4). ■

Lemma 4.83 (Local Estimate for the Mesh Cell Residual) *Under the assumption of Theorem 4.79, it is*

$$\begin{aligned} h_K \left\| \boldsymbol{r}_K^h \left(\boldsymbol{u}^h, p^h\right) \right\|_{L^2(K)} \le C \Big(&\left\| \nabla (\boldsymbol{u} - \boldsymbol{u}^h) \right\|_{L^2(K)} + \left\| p - p^h \right\|_{L^2(K)} \\ &+ h_K \left\| \boldsymbol{f} - P^h \boldsymbol{f} \right\|_{L^2(K)} \Big), \end{aligned} \tag{4.114}$$

with C independent of the mesh width and the solution.

Proof Considering a vector-valued version of (4.110) and choosing $\boldsymbol{r}_K^h \left(\boldsymbol{u}^h, p^h\right)$ as polynomial, which will be abbreviated in the proof by $\boldsymbol{r}_K^h$, gives

$$\left\| \boldsymbol{r}_K^h \right\|_{L^2(K)}^2 \le C \left(\boldsymbol{r}_K^h, \phi_K^h \boldsymbol{r}_K^h\right)_K. \tag{4.115}$$

Now, $\phi_K^h \boldsymbol{r}_K^h$ can be extended off K by zero which gives a function in V and which will be denoted by the same symbol. Using $\left(\phi_K^h \boldsymbol{r}_K^h, 0\right)$ as test function in (4.2) gives

$$\left(\nabla \boldsymbol{u}, \nabla\left(\phi_K^h \boldsymbol{r}_K^h\right)\right) - \left(\nabla \cdot \left(\phi_K^h \boldsymbol{r}_K^h\right), p\right) = \left(\boldsymbol{f}, \phi_K^h \boldsymbol{r}_K^h\right).$$

Adding and subtracting terms and applying integration by parts, using the definition (4.105) of the mesh cell residual, and observing that $\phi_K^h \boldsymbol{r}_K^h$ vanishes on ∂K yields

$$\begin{aligned}
&\left(\nabla(\boldsymbol{u} - \boldsymbol{u}^h), \nabla(\phi_K^h \boldsymbol{r}_K^h)\right) - \left(\nabla \cdot \left(\phi_K^h \boldsymbol{r}_K^h\right), p - p^h\right) \\
&= \left(P^h \boldsymbol{f}, \phi_K^h \boldsymbol{r}_K^h\right) + \left(\boldsymbol{f} - P^h \boldsymbol{f}, \phi_K^h \boldsymbol{r}_K^h\right) - \left(\nabla \boldsymbol{u}^h, \nabla(\phi_K^h \boldsymbol{r}_K^h\right) + \left(\nabla \cdot \left(\phi_K^h \boldsymbol{r}_K^h\right), p^h\right) \\
&= (\boldsymbol{r}_K^h, \phi_K^h \boldsymbol{r}_K^h)_K + \left(\boldsymbol{f} - P^h \boldsymbol{f}, \phi_K^h \boldsymbol{r}_K^h\right)_K.
\end{aligned}$$

Inserting this identity in (4.115), applying the Cauchy–Schwarz inequality (A.10), (3.41), and (4.111) leads to

$$\begin{aligned}
&\left\| \boldsymbol{r}_K^h \right\|_{L^2(K)}^2 \\
&\le C\Big(\left\| \nabla(\boldsymbol{u} - \boldsymbol{u}^h) \right\|_{L^2(K)} \left\| \nabla(\phi_K^h \boldsymbol{r}_K^h) \right\|_{L^2(K)} + \left\| p - p^h \right\|_{L^2(K)} \left\| \nabla(\phi_K^h \boldsymbol{r}_K^h) \right\|_{L^2(K)} \\
&\quad + \left\| \boldsymbol{f} - P^h \boldsymbol{f} \right\|_{L^2(K)} \left\| \phi_K^h \boldsymbol{r}_K^h \right\|_{L^2(K)} \Big) \\
&\le C\left(h_K^{-1} \left\| \nabla(\boldsymbol{u} - \boldsymbol{u}^h) \right\|_{L^2(K)} + h_K^{-1} \left\| p - p^h \right\|_{L^2(K)} + \left\| \boldsymbol{f} - P^h \boldsymbol{f} \right\|_{L^2(K)} \right) \left\| \boldsymbol{r}_K^h \right\|_{L^2(K)}.
\end{aligned}$$

Dividing by $h_K^{-1} \left\| \boldsymbol{r}_K^h \right\|_{L^2(K)}$ gives the statement of the lemma. ■

Lemma 4.84 (Local Estimate for the Face Residual) *With the assumption of Theorem 4.79, it is*

$$\begin{aligned}
h_E^{1/2} \left\| \boldsymbol{r}_E^h\left(\boldsymbol{u}^h, p^h\right) \right\|_{L^2(E)} \le C\Big(&\left\| \nabla(\boldsymbol{u} - \boldsymbol{u}^h) \right\|_{L^2(\omega_E)} + \left\| p - p^h \right\|_{L^2(\omega_E)} \\
&+ h_E \left\| \boldsymbol{f} - P^h \boldsymbol{f} \right\|_{L^2(\omega_E)} \Big),
\end{aligned} \tag{4.116}$$

with C independent of the mesh width and the solution.

Proof Taking the face residual (4.106), abbreviating this function with $\boldsymbol{r}_E^h$, and using a vector-valued version of (4.112) gives

$$\left\| \boldsymbol{r}_E^h \right\|_{L^2(E)}^2 \le C\left(\boldsymbol{r}_E^h, \phi_E^h \boldsymbol{r}_E^h\right)_E. \tag{4.117}$$

This function $\phi_E^h \boldsymbol{r}_E^h$ can be extended to a function in V, which is denoted with the same symbol, by setting $\phi_E^h \boldsymbol{r}_E^h$ outside ω_E to zero. Now, the test function $\left(\phi_E^h \boldsymbol{r}_E^h, 0\right)$

is applied in the Stokes equations (4.2) yielding

$$\left(\nabla \boldsymbol{u}, \nabla\left(\phi_E^h \boldsymbol{r}_E^h\right)\right) - \left(\nabla \cdot \left(\phi_E^h \boldsymbol{r}_E^h\right), p\right) = \left(\boldsymbol{f}, \phi_E^h \boldsymbol{r}_E^h\right).$$

Adding and subtracting terms and applying integration by parts, noting that $\phi_E^h \boldsymbol{r}_E^h$ vanishes on the boundary of ω_E and outside ω_E leads to

$$\begin{aligned}
&\left(\nabla(\boldsymbol{u}-\boldsymbol{u}^h), \nabla(\phi_E^h \boldsymbol{r}_E^h)\right) - \left(\nabla \cdot \left(\phi_E^h \boldsymbol{r}_E^h\right), p - p^h\right) \\
&= \left(P^h \boldsymbol{f}, \phi_E^h \boldsymbol{r}_E^h\right) + \left(\boldsymbol{f} - P^h \boldsymbol{f}, \phi_E^h \boldsymbol{r}_E^h\right) - \left(\nabla \boldsymbol{u}^h, \nabla\left(\phi_E^h \boldsymbol{r}_E^h\right)\right) + \left(\nabla \cdot \left(\phi_E^h \boldsymbol{r}_E^h\right), p^h\right) \\
&= \sum_{K \in \omega_E} \left[\left(\boldsymbol{r}_K^h, \phi_E^h \boldsymbol{r}_E^h\right)_K + \left(\boldsymbol{f} - P^h \boldsymbol{f}, \phi_E^h \boldsymbol{r}_E^h\right)_K\right] + \left(\boldsymbol{r}_E^h, \phi_E^h \boldsymbol{r}_E^h\right)_E .
\end{aligned}$$

This identity is inserted in (4.117). With the Cauchy–Schwarz inequality (A.10), estimate (3.41), and (4.113), one obtains

$$\begin{aligned}
&\left\|\boldsymbol{r}_E^h\right\|_{L^2(E)}^2 \\
&\le C\Big(\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\omega_E)} \left\|\nabla(\phi_E^h \boldsymbol{r}_E^h)\right\|_{L^2(\omega_E)} + \left\|p - p^h\right\|_{L^2(\omega_E)} \left\|\nabla(\phi_E^h \boldsymbol{r}_E^h)\right\|_{L^2(\omega_E)} \\
&\quad + \left\|\boldsymbol{r}_K^h\right\|_{L^2(\omega_E)} \left\|\phi_E^h \boldsymbol{r}_E^h\right\|_{L^2(\omega_E)} + \left\|\boldsymbol{f} - P^h \boldsymbol{f}\right\|_{L^2(\omega_E)} \left\|\phi_E^h \boldsymbol{r}_E^h\right\|_{L^2(\omega_E)} \Big) \\
&\le C\Big(h_E^{-1/2} \left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\omega_E)} + h_E^{-1/2} \left\|p - p^h\right\|_{L^2(\omega_E)} + h_E^{1/2} \left\|\boldsymbol{r}_K^h\right\|_{L^2(\omega_E)} \\
&\quad + h_E^{1/2} \left\|\boldsymbol{f} - P^h \boldsymbol{f}\right\|_{L^2(\omega_E)} \Big) \left\|\boldsymbol{r}_E^h\right\|_{L^2(\omega_E)} .
\end{aligned}$$

From the quasi-uniformity of the triangulation it follows that for each mesh cell K its diameter h_K can be estimated from below and above by a constant times h_E, where the constant is independent of the triangulation, of the concrete mesh cells, and of the edges. Using this equivalence, dividing by $h_E^{-1/2} \left\|\boldsymbol{r}_E^h\right\|_{L^2(\omega_E)}$, and inserting estimate (4.116) gives the estimate for the face residual. ■

Theorem 4.85 (Local Lower, Residual-Based, a Posteriori Error Estimate for Conforming Inf-Sup Stable Finite Element Spaces) *Let the assumptions of Theorem 4.79 be satisfied and let*

$$\omega_K = \bigcup_{K' \in \mathcal{T}^h, \mathcal{E}(K') \subset \mathcal{E}(K) \neq \emptyset} K',$$

then there holds the estimate

$$\eta_K \le C\left(\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\omega_K)} + \left\|p - p^h\right\|_{L^2(\omega_K)} + h_K \left\|\boldsymbol{f} - P^h \boldsymbol{f}\right\|_{L^2(\omega_K)}\right), \qquad (4.118)$$

where the constant is independent of the mesh width and of the solution of the Stokes equations.

Proof It is with (3.40)

$$\left\| \nabla \cdot \boldsymbol{u}^h \right\|_{L^2(K)} = \left\| \nabla \cdot \left(\boldsymbol{u} - \boldsymbol{u}^h \right) \right\|_{L^2(K)} \le \sqrt{d} \left\| \nabla \left(\boldsymbol{u} - \boldsymbol{u}^h \right) \right\|_{L^2(K)}.$$

With this estimate, estimate (4.114), and estimate (4.116) for all faces of K, (4.118) follows immediately. ■

Remark 4.86 (On the Local Lower Estimate (4.118)*)*

- The use of residual-based error estimators for controlling an adaptive mesh refinement is popular in academia. Besides theoretical support, like estimates (4.118), the rather easy implementation of such estimators is a main reason.
- The constant in estimate (4.118) comes from the estimates for the local bubble functions given in Lemma 4.82.

□

Example 4.87 (Global a Posteriori Error Estimation) Exactly the same problem and the same setup as in Example 4.31 is considered. Comparisons of the actual errors and the error estimates for two discretizations are presented in Fig. 4.11. For both discretizations, the error is dominated by $\left\| \nabla (\boldsymbol{u} - \boldsymbol{u}^h) \right\|_{L^2(\Omega)}$, compare Figs. 4.3 and 4.5. It can be seen that in both cases the error and the estimator possess the same order of convergence. However, the estimator overestimates the error by a factor of around 13 for P_2/P_1 and of around 20 for P_3/P_2. These factors represent the unknown constant C in the estimate (4.108). □

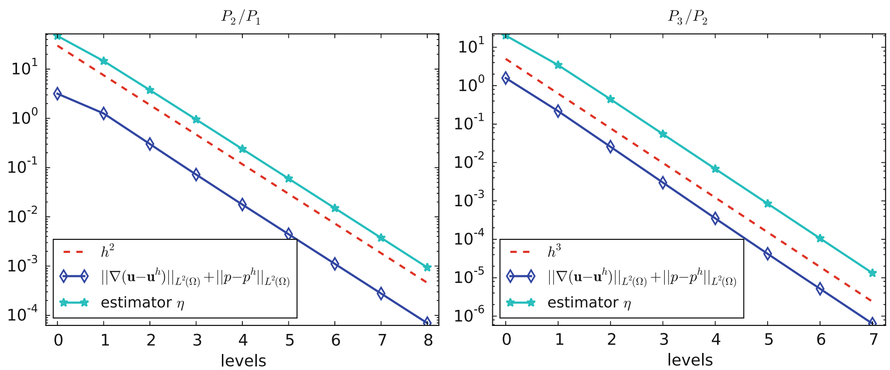

Fig. 4.11 Example 4.87. Convergence of the errors $\left\| \nabla (\boldsymbol{u} - \boldsymbol{u}^h) \right\|_{L^2(\Omega)} + \left\| p - p^h \right\|_{L^2(\Omega)}$ and of the error estimator $\eta = \left(\sum_{K \in \mathcal{T}^h} \eta_K^2 \right)^{1/2}$ with η_K from (4.107) for the pairs P_2/P_1 and P_3/P_2

Example 4.88 (Adaptive Grid Refinement) This examples considers the setup of the regularized driven cavity problem, see Example D.4 for the Stokes equations. The solution obtained for the Stokes equations looks similar to the solution for the Navier–Stokes equations with $\nu = 1$, compare Fig. D.2. Thus, the main changes of the solution occur at the upper two corners of the domain. The change of the velocity is imposed by the prescribed boundary conditions and the pressure shows a peak in each corner.

The adaptive grid refinement should start on a grid where the most essential features of the solution are already visible. Figure 4.12 presents a typical approach for determining the adaptively refined grid. Usually, all mesh cells K are marked for refinement with

$$\eta_K \geq C_0 \overline{\eta},$$

where $\overline{\eta}$ is some reference value. For the simulation presented in this example,

$$\overline{\eta} = \max_{K \in \mathcal{T}^h} \eta_K$$

and $C_0 = 0.5$ were chosen. In the case of stationary problems, it is important for obtaining an efficient algorithm that there is a sufficient increase of the number of degrees of freedom after an adaptive refinement. In this way, one prevents the expensive assembling and solving of a discrete system with only few new degrees of freedom, which generally results only in a minor improvement of the discrete solution. In the simulation presented here, it was required that at least 10 % of the mesh cells were marked for refinement.

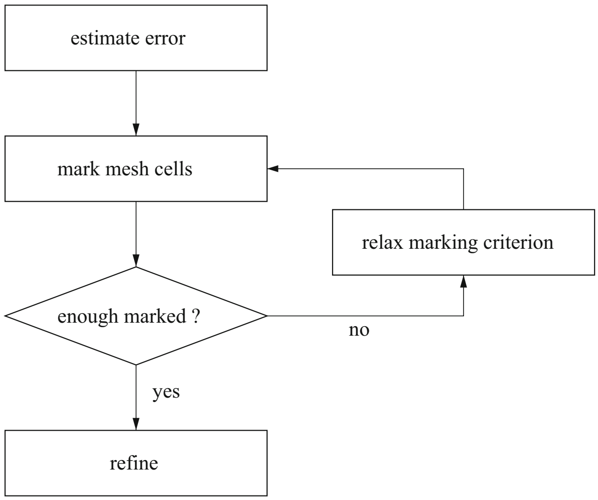

Fig. 4.12 Example 4.87. Algorithm for determining the adaptively refined grid

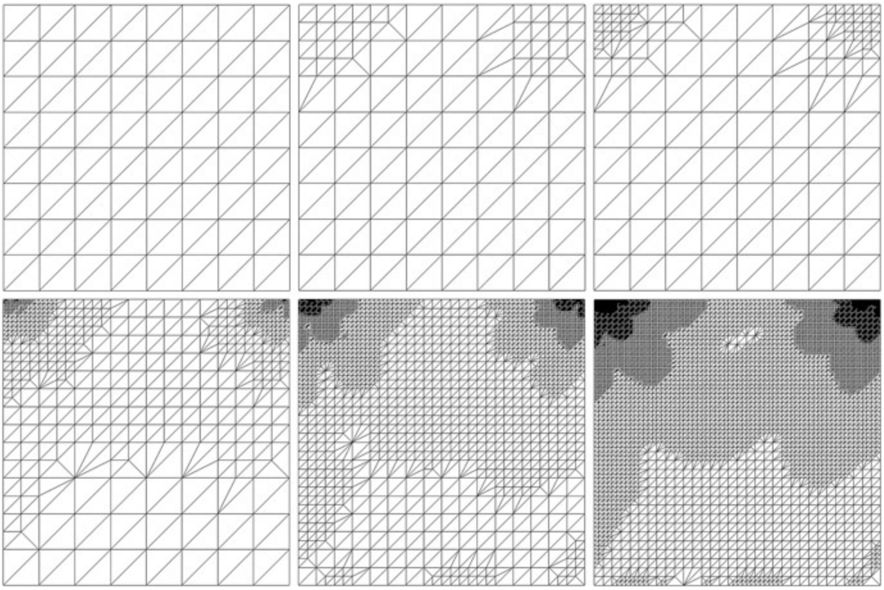

Fig. 4.13 Example 4.88. Adaptive grid refinement based on the local error estimators $\{\eta_K\}$ defined in (4.107) for the regularized driven cavity problem

An adaptive grid refinement for a simulation with the P_2/P_1 pair of finite element spaces is presented in Fig. 4.13. Since the most complex part of the solution is in the upper two corners, one expects that the adaptive grid refinement should start there. This expectation is met. Next, the subregion with the big vortex is refined and finally there is some refinement in the lower part of the domain where almost no flow occurs. □

Remark 4.89 (A Posteriori Error Analysis for the Crouzeix–Raviart Pair of Spaces P_1^{nc}/P_0) The derivation of residual-based a posteriori error estimators for the P_1^{nc}/P_0 pair of finite element spaces for the Stokes equations in a simply connected two-dimensional domain is presented in Dari et al. (1995). An additional tool compared with the conforming case is an appropriate decomposition of the error $\sum_{K\in\mathcal{T}^h} \nabla\left(\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\big|_K\right)$. This decomposition is similar to the Helmholtz decomposition, see Theorem 3.168. □

4.5 Stabilized Finite Element Methods Circumventing the Discrete Inf-Sup Condition

Remark 4.90 (Motivation) The application of inf-sup stable pairs of finite elements requires the use of different spaces for velocity and pressure. Moreover, it is not possible to use conforming spaces of lowest order for the discrete velocity, i.e.,

P_1 or Q_1 finite element spaces. However, software for solving incompressible flow problems contains often just one finite element space and then usually P_1 or Q_1. In such situations, it is necessary to modify the discrete problem such that the satisfaction of the discrete inf-sup condition (3.51) is not longer necessary. To this end, one has to remove the saddle point structure of the discrete problem, i.e., one has to remove the zero matrix block in the pressure-pressure coupling of the saddle point problem.

There are several proposals in the literature for defining appropriate pressure-pressure couplings, so-called stabilization terms. One class are so-called pseudo-compressibility methods. The strong form of such methods might look as follows:

- Pressure Stabilization Petrov–Galerkin (PSPG) method

$$-\nabla \cdot \boldsymbol{u} + \delta \Delta p = 0,$$

- penalty method

$$-\nabla \cdot \boldsymbol{u} - \delta p = 0,$$

- artificial compressibility method

$$-\nabla \cdot \boldsymbol{u} - \delta \partial_t p = 0,$$

where δ is some parameter which has to be chosen appropriately.

Note that the introduction of a pressure-pressure coupling perturbs the continuity equation and thus the conservation of mass. In addition, each stabilization term contains parameters. The asymptotic optimal choice of stabilization parameters can be determined often with results from numerical analysis, e.g., with conditions for the existence and uniqueness of a solution of the stabilized problem or from optimal error estimates. However, the user still has to choose concrete parameters in simulations and, depending on the parameter, different concrete choices of the same asymptotic type might sometimes lead to rather different numerical solutions. □

Remark 4.91 (Contents of this Section) The probably most popular approach which circumvents the discrete inf-sup condition is the PSPG method. This method is presented and discussed in detail in Sect. 4.5.1. Some alternative approaches are sketched in Sect. 4.5.2. □

4.5.1 The Pressure Stabilization Petrov–Galerkin (PSPG) Method

Remark 4.92 (The Pressure Stabilization Petrov–Galerkin (PSPG) Method) The PSPG method was proposed for finite element spaces with continuous discrete pressures in Hughes et al. (1986). In the case of piecewise polynomial but discontinuous finite element pressure spaces, an additional term is necessary, which was introduced in Hughes and Franca (1987), Douglas and Wang (1989).

The PSPG method has the form: Given $\boldsymbol{f} \in L^2(\Omega)$, find $\left(\boldsymbol{u}^h, p^h\right) \in V^h \times Q^h$ such that

$$A_{\text{pspg}}\left(\left(\boldsymbol{u}^h, p^h\right), \left(\boldsymbol{v}^h, q^h\right)\right) = L_{\text{pspg}}\left(\left(\boldsymbol{v}^h, q^h\right)\right) \quad \forall \left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h, \tag{4.119}$$

where the bilinear form $A_{\text{pspg}} : \left(V \times \tilde{Q}\right) \times \left(V \times \tilde{Q}\right) \to \mathbb{R}$ is given by

$$\begin{aligned} & A_{\text{pspg}}\left((\boldsymbol{u}, p), (\boldsymbol{v}, q)\right) \\ &= \nu\left(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}\right) - \left(\nabla \cdot \boldsymbol{v}, p\right) + \left(\nabla \cdot \boldsymbol{u}, q\right) \\ &\quad + \sum_{E \in \mathcal{E}^h} \gamma_E \left([|p|]_E, [|q|]_E\right)_E + \sum_{K \in \mathcal{T}^h} \left(-\nu \Delta \boldsymbol{u} + \nabla p, \delta_K^p \nabla q\right)_K \end{aligned} \tag{4.120}$$

and the linear form $L_{\text{pspg}} : V \times \tilde{Q} \to \mathbb{R}$ by

$$L_{\text{pspg}}\left((\boldsymbol{v}, q)\right) = (\boldsymbol{f}, \boldsymbol{v}) + \sum_{K \in \mathcal{T}^h} \left(\boldsymbol{f}, \delta_K^p \nabla q\right)_K, \tag{4.121}$$

with

$$\tilde{Q} = Q \cap \left\{q \in Q : q|_K \in H^1(K) \text{ for all } K \in \mathcal{T}^h\right\}. \tag{4.122}$$

The definition of $\tilde{Q}$ ensures that the jumps of the pressure across the faces of the mesh cells are well defined. If $Q^h \subset H^1(\Omega)$, then the jumps of the pressure vanish almost everywhere on the faces. From the practical point of view, the case of piecewise polynomial and continuous discrete pressure functions is very important such that then even $Q^h \subset C(\overline{\Omega})$.

The volume integral in the stabilization term in (4.120) contains the so-called strong residual of the Stokes equations. The basic idea behind the use of the strong residual will be discussed in Remark 5.21.

Because of the stabilization terms, problem (4.119) is not a saddle point problem. Appropriate choices for the stabilization parameter $\{\delta_K^p\}$ and $\{\gamma_E\}$ will be based on the study of the existence and uniqueness of a solution of (4.119), see Lemma 4.95, and on finite element error estimates, see Theorem 4.98. □

Remark 4.93 (Artificial Pressure Boundary Conditions) Consider for simplicity of presentation the situation that $Q^h \subset H^1(\Omega)$. The additional pressure-pressure coupling in (4.120) represents the discrete formulation of a Laplacian operator. Thus, the PSPG method introduces a (discretization of a) second order partial differential operator for the pressure. To obtain a well-posed problem, there have to be appropriate boundary conditions for the corresponding continuous pressure, i.e., one needs boundary conditions for the pressure on the whole boundary. Since the physical modeling of the Stokes and Navier–Stokes equations does not provide such boundary conditions, these conditions are artificial.

Since the pressure space Q^h for the PSPG method does not contain essential boundary conditions, the artificial boundary conditions are natural ones, i.e., they are Neumann boundary conditions. These conditions can be derived by assuming that all functions in the Stokes equations are sufficiently smooth and then by applying the negative of the divergence operator to these equations. One gets with $\nabla \cdot \boldsymbol{u} = 0$ and

$$\begin{aligned}\nabla \cdot \Delta \boldsymbol{u} &= \partial_x \left(\partial_{xx} u_1 + \partial_{yy} u_1 + \partial_{zz} u_1\right) + \partial_y \left(\partial_{xx} u_2 + \partial_{yy} u_2 + \partial_{zz} u_2\right) \\ &\quad + \partial_z \left(\partial_{xx} u_3 + \partial_{yy} u_3 + \partial_{zz} u_3\right) \\ &= \partial_{xx} \left(\partial_x u_1 + \partial_y u_2 + \partial_z u_3\right) + \partial_{xx} \left(\partial_x u_1 + \partial_y u_2 + \partial_z u_3\right) \\ &\quad + \partial_{xx} \left(\partial_x u_1 + \partial_y u_2 + \partial_z u_3\right) \\ &= 0, \end{aligned} \tag{4.123}$$

that

$$-\Delta p = -\nabla \cdot \boldsymbol{f} \quad \text{in } \Omega.$$

Thus, one finds, by testing with a function $q \in Q$, where for the moment $q \in H^1(\Omega)$ will be assumed, and integrating by parts that

$$(\nabla p, \nabla q) - \int_\Gamma (\nabla p \cdot \boldsymbol{n})\, q(\boldsymbol{s})\, d\boldsymbol{s} = (\boldsymbol{f}, \nabla q) - \int_\Gamma (\boldsymbol{f} \cdot \boldsymbol{n})\, q(\boldsymbol{s})\, d\boldsymbol{s}.$$

Comparing this formulation with (4.120) and (4.121), one obtains that in the case $\delta_K^p = \delta$ for all $K \in \mathcal{T}^h$ the Neumann boundary conditions are of the form

$$\delta \left(\nabla p - \boldsymbol{f}\right) \cdot \boldsymbol{n} = 0 \quad \text{on } \Gamma.$$

□

Lemma 4.94 (A Norm in $V^h \times Q^h$ Containing the Stabilization Terms) *Let $\delta_K^p > 0$ for all $K \in \mathcal{T}^h$, in the case $Q^h \not\subset H^1(\Omega)$ let $\gamma_E > 0$ for all $E \in \mathcal{E}^h$, and let V^h and Q^h be conforming finite element spaces. Then,*

$$\left\| \left(\boldsymbol{v}^h, q^h\right) \right\|_{\text{pspg}} = \left(\nu \left\| \nabla \boldsymbol{v}^h \right\|_{L^2(\Omega)}^2 + \sum_{E \in \mathcal{E}^h} \gamma_E \left\| \left[\left| q^h \right|\right]_E \right\|_{L^2(E)}^2 + \sum_{K \in \mathcal{T}^h} \delta_K^p \left\| \nabla q^h \right\|_{L^2(K)}^2 \right)^{1/2} \tag{4.124}$$

defines a norm in $V^h \times Q^h$.

Proof Expression (4.124) is the square root of a sum of squares of seminorms. Thus, it is clearly a seminorm itself. It remains to prove that from $\left\|\left(\boldsymbol{v}^h, q^h\right)\right\|_{\text{pspg}} = 0$ it follows that $\boldsymbol{v}^h = \boldsymbol{0}$ and $q^h = 0$, see Definition A.6.

Let $\left\|\left(\boldsymbol{v}^h, q^h\right)\right\|_{\text{pspg}} = 0$, then all terms in (4.124) vanish. In particular, it holds $\left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)} = 0$. Since this expression is a norm in V^h, it follows that $\boldsymbol{v}^h = \boldsymbol{0}$. With this result, one gets

$$0 = \sum_{E\in\mathcal{E}^h} \gamma_E \left\|[|q|]_E\right\|_{L^2(E)}^2 + \sum_{K\in\mathcal{T}^h} \delta_K^p \left\|\nabla q^h\right\|_{L^2(K)}^2 .$$

Because δ_K^p is assumed to be positive for all mesh cells, it follows that $\left\|\nabla q^h\right\|_{L^2(K)} = 0$ for all $K \in \mathcal{T}^h$. If $Q^h \subset H^1(\Omega)$, then $\left\|[|q|]_E\right\|_{L^2(E)} = 0$ for all faces. Otherwise, one gets this property from the assumption $\gamma_E > 0$ for all faces. Altogether, it follows that q^h is constant on Ω. The only globally constant function in Q^h is $q^h = 0$. Hence $\left\|\left(\boldsymbol{v}^h, q^h\right)\right\|_{\text{pspg}}$ defines a norm on $V^h \times Q^h$. ■

Lemma 4.95 (Existence and Uniqueness of a Solution, Stability Estimate) *Let the assumption of Lemma 4.94 be satisfied and let*

$$\delta_K^p \le \frac{h_K^2}{\nu C_{\text{inv}}^2}, \tag{4.125}$$

then the PSPG problem (4.119) possesses a unique solution. This solution satisfies the stability estimate

$$\left\|\left(\boldsymbol{u}^h, p^h\right)\right\|_{\text{pspg}} \le 2\sqrt{2}\left(\frac{C}{\nu^{1/2}}\left\|\boldsymbol{f}\right\|_{L^2(\Omega)} + \left(\sum_{K\in\mathcal{T}^h} \delta_K^p \left\|\boldsymbol{f}\right\|_{L^2(K)}^2\right)^{1/2}\right). \tag{4.126}$$

Proof Since the bilinear problem (4.119) is not of saddle point type, the Theorem of Lax–Milgram, Theorem B.4, can be used for proving the existence and uniqueness of a solution.

First, the coercivity of the bilinear form $A_{\text{pspg}}(\cdot,\cdot)$ with respect to the norm $\|\cdot\|_{\text{pspg}}$ will be shown. One obtains with the Cauchy–Schwarz inequality (A.10), the definition (4.124) of the norm, the inverse inequality (C.35), Young's inequality (A.5), and the condition (4.125) on the stabilization parameters

$$\begin{aligned}
& A_{\text{pspg}}\left(\left(\boldsymbol{v}^h, q^h\right), \left(\boldsymbol{v}^h, q^h\right)\right) \\
& \ge \nu \left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)}^2 + \sum_{E\in\mathcal{E}^h} \gamma_E \left\|\left[\left|q^h\right|\right]_E\right\|_{L^2(E)}^2 + \sum_{K\in\mathcal{T}^h} \delta_K^p \left\|\nabla q^h\right\|_{L^2(K)}^2 \\
& \quad - \sum_{K\in\mathcal{T}^h} \delta_K^p \nu \left\|\Delta \boldsymbol{v}^h\right\|_{L^2(K)} \left\|\nabla q^h\right\|_{L^2(K)}
\end{aligned}$$

$$
\begin{aligned}
&\geq \left\| \left(\boldsymbol{v}^h, q^h\right) \right\|_{\text{pspg}}^2 - \sum_{K \in \mathcal{T}^h} \frac{\delta_K^p C_{\text{inv}} \nu}{h_K} \left\| \nabla \boldsymbol{v}^h \right\|_{L^2(K)} \left\| \nabla q^h \right\|_{L^2(K)} \\
&\geq \left\| \left(\boldsymbol{v}^h, q^h\right) \right\|_{\text{pspg}}^2 - \sum_{K \in \mathcal{T}^h} \frac{1}{2} \frac{\delta_K^p C_{\text{inv}}^2 \nu^2}{h_K^2} \left\| \nabla \boldsymbol{v}^h \right\|_{L^2(K)}^2 - \frac{1}{2} \sum_{K \in \mathcal{T}^h} \delta_K^p \left\| \nabla q^h \right\|_{L^2(K)}^2 \\
&\geq \frac{1}{2} \left\| \left(\boldsymbol{v}^h, q^h\right) \right\|_{\text{pspg}}^2 ,
\end{aligned}
\tag{4.127}
$$

for all $\left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h$. The bilinear form A_{pspg} is bounded since it is continuous. All terms of A_{pspg} are defined with integrals and the continuity follows from the continuity of the integrals.

The boundedness of the right-hand side of (4.119) will be established with a direct estimation, using the Cauchy–Schwarz inequality, Poincaré's inequality (A.12), the Cauchy–Schwarz inequality for sums (A.2), the inequality $ab + cd \leq (a+c)(b+d)$ which is valid for non-negative real numbers a, b, c, d, the estimate $a + b \leq \sqrt{2}(a^2 + b^2)^{1/2}$, which is a consequence of Young's inequality, and the definition (4.124) of the norm

$$
\begin{aligned}
&(\boldsymbol{f}, \boldsymbol{v}^h) + \sum_{K \in \mathcal{T}^h} \left(\boldsymbol{f}, \delta_K^p \nabla q^h\right)_K \\
&\leq \|\boldsymbol{f}\|_{L^2(\Omega)} \left\| \boldsymbol{v}^h \right\|_{L^2(\Omega)} + \sum_{K \in \mathcal{T}^h} \delta_K^p \|\boldsymbol{f}\|_{L^2(K)} \left\| \nabla q^h \right\|_{L^2(K)} \\
&\leq C \|\boldsymbol{f}\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{v}^h \right\|_{L^2(\Omega)} + \left(\sum_{K \in \mathcal{T}^h} \delta_K^p \|\boldsymbol{f}\|_{L^2(K)}^2 \right)^{1/2} \left(\sum_{K \in \mathcal{T}^h} \delta_K^p \left\| \nabla q^h \right\|_{L^2(K)}^2 \right)^{1/2} \\
&\leq \left(\frac{C}{\nu^{1/2}} \|\boldsymbol{f}\|_{L^2(\Omega)} + \left(\sum_{K \in \mathcal{T}^h} \delta_K^p \|\boldsymbol{f}\|_{L^2(K)}^2 \right)^{1/2} \right) \\
&\quad \times \left(\nu^{1/2} \left\| \nabla \boldsymbol{v}^h \right\|_{L^2(\Omega)} + \left(\sum_{K \in \mathcal{T}^h} \delta_K^p \left\| \nabla q^h \right\|_{L^2(K)}^2 \right)^{1/2} \right) \\
&\leq \sqrt{2} \left(\frac{C}{\nu^{1/2}} \|\boldsymbol{f}\|_{L^2(\Omega)} + \left(\sum_{K \in \mathcal{T}^h} \delta_K^p \|\boldsymbol{f}\|_{L^2(K)}^2 \right)^{1/2} \right) \\
&\quad \times \left(\nu \left\| \nabla \boldsymbol{v}^h \right\|_{L^2(\Omega)}^2 + \sum_{K \in \mathcal{T}^h} \delta_K^p \left\| \nabla q^h \right\|_{L^2(K)}^2 \right)^{1/2} \\
&\leq \sqrt{2} \left(\frac{C}{\nu^{1/2}} \|\boldsymbol{f}\|_{L^2(\Omega)} + \left(\sum_{K \in \mathcal{T}^h} \delta_K^p \|\boldsymbol{f}\|_{L^2(K)}^2 \right)^{1/2} \right) \left\| \left(\boldsymbol{v}^h, q^h\right) \right\|_{\text{pspg}} ,
\end{aligned}
\tag{4.128}
$$

for all $\left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h$. Thus, the conditions of the Theorem of Lax–Milgram are satisfied from which it follows that (4.119) possesses a unique solution.

The stability estimate (4.126) follows from (4.127) and (4.119), yielding

$$\frac{1}{2}\left\|\left(\boldsymbol{u}^h, p^h\right)\right\|_{\mathrm{pspg}}^2 \leq A_{\mathrm{pspg}}\left(\left(\boldsymbol{u}^h, p^h\right), \left(\boldsymbol{u}^h, p^h\right)\right) = L_{\mathrm{pspg}}\left(\left(\boldsymbol{u}^h, p^h\right)\right)$$

and from inserting (4.128) in the right-hand side of this inequality. ■

Remark 4.96 ($\|\cdot\|_{\mathrm{pspg}}$ is a Mesh-Dependent Norm) Since the stabilization parameters have to satisfy (4.125), they depend on the local mesh size. Hence, the norm $\|\cdot\|_{\mathrm{pspg}}$ is a mesh-dependent norm. □

Lemma 4.97 (Galerkin Orthogonality) *Let $(\boldsymbol{u}, p)$ be the solution of (4.44), which is assumed to be sufficiently smooth, and let $\left(\boldsymbol{u}^h, p^h\right) \in V^h \times Q^h$ be the solution of the PSPG problem (4.119), then*

$$A_{\mathrm{pspg}}\left(\left(\boldsymbol{u}-\boldsymbol{u}^h, p-p^h\right), \left(\boldsymbol{v}^h, q^h\right)\right) = 0 \quad \forall\ \left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h. \tag{4.129}$$

Proof Using the linearity of the bilinear form, (4.119), and the smoothness assumption on the solution of (4.44) yields

$$\begin{aligned}
& A_{\mathrm{pspg}}\left(\left(\boldsymbol{u}-\boldsymbol{u}^h, p-p^h\right), \left(\boldsymbol{v}^h, q^h\right)\right) \\
&= A_{\mathrm{pspg}}\left((\boldsymbol{u}, p), \left(\boldsymbol{v}^h, q^h\right)\right) - A_{\mathrm{pspg}}\left(\left(\boldsymbol{u}^h, p^h\right), \left(\boldsymbol{v}^h, q^h\right)\right) \\
&= \nu\,(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) - (\nabla\cdot\boldsymbol{v}, p) + (\nabla\cdot\boldsymbol{u}, q) + \sum_{K\in\mathcal{T}^h}\left(-\nu\Delta\boldsymbol{u} + \nabla p, \delta_K^p \nabla q^h\right)_K \\
&\quad - L_{\mathrm{pspg}}\left(\left(\boldsymbol{v}^h, q^h\right)\right) \\
&= (\boldsymbol{f}, \boldsymbol{v}) + \sum_{K\in\mathcal{T}^h}\left(\boldsymbol{f}, \delta_K^p \nabla q^h\right)_K - L_{\mathrm{pspg}}\left(\left(\boldsymbol{v}^h, q^h\right)\right) = 0.
\end{aligned}$$

■

Theorem 4.98 (Finite Element Error Estimate) *Let $(\boldsymbol{u}, p)$ be the solution of (4.44) and let $\left(\boldsymbol{u}^h, p^h\right) \in V^h \times Q^h$ be the solution of the PSPG problem (4.119). Assume that a quasi-uniform family of triangulations is used,*

$$P_k \text{ or } Q_k \subseteq V^h \subset V, \quad k \geq 1, \qquad P_l \text{ or } Q_l \subseteq Q^h \subset Q, \quad l \geq 0,$$

assume that $\boldsymbol{u} \in H^{k+1}(\Omega)$, $p \in H^{l+1}(\Omega)$, and that the stabilization parameters are chosen as follows

$$\delta_K^p = C_0 \frac{h_K^2}{\nu}, \quad \gamma_E = C_1 \frac{h_E}{\nu}, \tag{4.130}$$

with C_0 such that (4.125) is satisfied. Then, the following error estimate holds

$$\left\| \left(\boldsymbol{u} - \boldsymbol{u}^h, p - p^h\right) \right\|_{\text{pspg}} \le C \left(\nu^{1/2} h^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \frac{h^{l+1}}{\nu^{1/2}} \|p\|_{H^{l+1}(\Omega)} \right). \tag{4.131}$$

Proof The triangle inequality gives

$$\left\| \left(\boldsymbol{u} - \boldsymbol{u}^h, p - p^h\right) \right\|_{\text{pspg}} \le \left\| \left(\boldsymbol{u} - I^h\boldsymbol{u}, p - P^h_{L^2}p\right) \right\|_{\text{pspg}} + \left\| \left(\boldsymbol{u}^h - I^h\boldsymbol{u}, p^h - P^h_{L^2}p\right) \right\|_{\text{pspg}}, \tag{4.132}$$

where $I^h\boldsymbol{u} \in V^h$ is the Lagrangian interpolant of $\boldsymbol{u}$ and $P^h_{L^2}p \in P_l$ (or Q_l) is the $L^2(\Omega)$ projection. By assumption on the pressure finite element space, it is $P^h_{L^2}p \in Q^h$. Both terms on the right-hand side of (4.132) are estimated separately.

For the interpolation error, one notes that the term with the pressure jumps vanishes for $l \ge 1$. One obtains with the interpolation estimates (C.14), (C.34) for $l = 0$, and (C.29), and with the assumptions (4.130) on the stabilization parameters

$$\begin{aligned}
&\left\| \left(\boldsymbol{u} - I^h\boldsymbol{u}, p - P^h_{L^2}p\right) \right\|^2_{\text{pspg}} \\
&= \nu \left\| \nabla \left(\boldsymbol{u} - I^h\boldsymbol{u}\right) \right\|^2_{L^2(\Omega)} + \sum_{E \in \mathcal{E}^h} \gamma_E \left\| \left[\left|p - P^h_{L^2}p\right|\right]_E \right\|^2_{L^2(E)} \\
&\quad + \sum_{K \in \mathcal{T}^h} \delta^p_K \left\| \nabla \left(p - P^h_{L^2}p\right) \right\|^2_{L^2(K)} \\
&\le \nu \left\| \nabla \left(\boldsymbol{u} - I^h\boldsymbol{u}\right) \right\|^2_{L^2(\Omega)} + \frac{Ch}{\nu} \sum_{E \in \mathcal{E}^h} \left\| \left[\left|p - P^h_{L^2}p\right|\right]_E \right\|^2_{L^2(E)} \\
&\quad + Ch^2 \left\| \nabla \left(p - P^h_{L^2}p\right) \right\|^2_{L^2(\Omega)} \\
&\le C \left(\nu h^{2k} \|\boldsymbol{u}\|^2_{H^{k+1}(\Omega)} + \frac{h^{2(l+1)}}{\nu} \|p\|^2_{H^{l+1}(\Omega)} + \frac{h^{2(l+1)}}{\nu} \|p\|^2_{H^{l+1}(\Omega)} \right) \\
&= C \left(\nu h^{2k} \|\boldsymbol{u}\|^2_{H^{k+1}(\Omega)} + \frac{h^{2(l+1)}}{\nu} \|p\|^2_{H^{l+1}(\Omega)} \right).
\end{aligned}$$

The estimate of the second term of (4.132) starts with the coercivity (4.127) and the Galerkin orthogonality (4.129)

$$\begin{aligned}
&\left\| \left(\boldsymbol{u}^h - I^h\boldsymbol{u}, p^h - P^h_{L^2}p\right) \right\|_{\text{pspg}} \\
&\le 2A_{\text{pspg}} \left(\left(\boldsymbol{u}^h - I^h\boldsymbol{u}, p^h - P^h_{L^2}p\right), \left(\boldsymbol{u}^h - I^h\boldsymbol{u}, p^h - P^h_{L^2}p\right) \right) \\
&= 2A_{\text{pspg}} \left(\left(\boldsymbol{u} - I^h\boldsymbol{u}, p - P^h_{L^2}p\right), \left(\boldsymbol{u}^h - I^h\boldsymbol{u}, p^h - P^h_{L^2}p\right) \right).
\end{aligned} \tag{4.133}$$

Now, each term of the right-hand side of (4.133) is estimated separately. The goal of these estimates is to obtain interpolation errors and to hide the other terms in the left-hand side of (4.133).

Using the Cauchy–Schwarz inequality (A.10), Young's inequality (A.5), and the interpolation estimate (C.14), one obtains for the viscous term

$$
\begin{aligned}
&\nu\left(\nabla\left(\boldsymbol{u}-I^h\boldsymbol{u}\right),\nabla\left(\boldsymbol{u}^h-I^h\boldsymbol{u}\right)\right)\\
&\le \nu\left\|\nabla\left(\boldsymbol{u}-I^h\boldsymbol{u}\right)\right\|_{L^2(\Omega)}\left\|\nabla\left(\boldsymbol{u}^h-I^h\boldsymbol{u}\right)\right\|_{L^2(\Omega)}\\
&\le 4\nu h^{2k}\left\|\nabla\left(\boldsymbol{u}-I^h\boldsymbol{u}\right)\right\|_{L^2(\Omega)}^2+\frac{\nu}{16}\left\|\nabla\left(\boldsymbol{u}^h-I^h\boldsymbol{u}\right)\right\|_{L^2(\Omega)}^2\\
&\le C\nu h^{2k}\left\|\boldsymbol{u}\right\|_{H^{k+1}(\Omega)}^2+\frac{\nu}{16}\left\|\nabla\left(\boldsymbol{u}^h-I^h\boldsymbol{u}\right)\right\|_{L^2(\Omega)}^2.
\end{aligned}
$$

The last term can be absorbed in the left-hand side of (4.133). In a similar way, using (3.170) and (C.28), one gets

$$
\left(\nabla\cdot\left(\boldsymbol{u}^h-I^h\boldsymbol{u}\right),p-P_{L^2}^h p\right)\le C\frac{h^{2(l+1)}}{\nu}\left\|p\right\|_{H^{l+1}(\Omega)}^2+\frac{\nu}{16}\left\|\nabla\left(\boldsymbol{u}^h-I^h\boldsymbol{u}\right)\right\|_{L^2(\Omega)}^2.
$$

The estimate of the next term requires an integration by parts

$$
\begin{aligned}
&\left(\nabla\cdot\left(\boldsymbol{u}-I^h\boldsymbol{u}\right),p^h-P_{L^2}^h p\right) \\
&=\sum_{E\in\mathcal{E}^h}\left(\left(\boldsymbol{u}-I^h\boldsymbol{u}\right)\cdot\boldsymbol{n}_E,\left[\left|p^h-P_{L^2}^h p\right|\right]_E\right)_E-\sum_{K\in\mathcal{T}^h}\left(\boldsymbol{u}-I^h\boldsymbol{u},\nabla\left(p^h-P_{L^2}^h p\right)\right)_K.
\end{aligned}
\tag{4.134}
$$

Both terms on the right-hand side of (4.134) are estimated more or less in the same way, e.g., one obtains for the last term with the Cauchy–Schwarz inequality, the Cauchy–Schwarz inequality for sums (A.2), Young's inequality, the definition (4.130) of the stabilization parameters, and the interpolation estimate (C.7)

$$
\begin{aligned}
&\sum_{K\in\mathcal{T}^h}\left(\boldsymbol{u}-I^h\boldsymbol{u},\nabla\left(p^h-P_{L^2}^h p\right)\right)_K\\
&\le\left(\sum_{K\in\mathcal{T}^h}\frac{1}{\delta_K^p}\left\|\boldsymbol{u}-I^h\boldsymbol{u}\right\|_{L^2(K)}^2\right)^{1/2}\left(\sum_{K\in\mathcal{T}^h}\delta_K^p\left\|\nabla\left(p^h-P_{L^2}^h p\right)\right\|_{L^2(K)}^2\right)^{1/2}\\
&\le C\sum_{K\in\mathcal{T}^h}\frac{\nu}{h_K^2}h_K^{2(k+1)}\left\|\boldsymbol{u}\right\|_{H^{k+1}(K)}^2+\frac{1}{16}\sum_{K\in\mathcal{T}^h}\delta_K^p\left\|\nabla\left(p^h-P_{L^2}^h p\right)\right\|_{L^2(K)}^2\\
&\le C\nu h^{2k}\left\|\boldsymbol{u}\right\|_{H^{k+1}(\Omega)}^2+\frac{1}{16}\sum_{K\in\mathcal{T}^h}\delta_K^p\left\|\nabla\left(p^h-P_{L^2}^h p\right)\right\|_{L^2(K)}^2.
\end{aligned}
$$

The estimate of the other term on the right-hand side of (4.134) uses (C.15). All stabilization terms are estimated with the same tools used so far. One gets

$$\begin{aligned}&\sum_{K\in\mathcal{T}^h}\left(-\nu\Delta\left(\boldsymbol{u}-I^h\boldsymbol{u}\right),\delta_K^p\nabla\left(p^h-P_{L^2}^h p\right)\right)_K\\&\le C\nu h^{2k}\left\|\boldsymbol{u}\right\|_{H^{k+1}(\Omega)}^2+\frac{1}{16}\sum_{K\in\mathcal{T}^h}\delta_K^p\left\|\nabla\left(p^h-P_{L^2}^h p\right)\right\|_{L^2(K)}^2,\end{aligned}$$

and

$$\begin{aligned}&\sum_{K\in\mathcal{T}^h}\left(\nabla\left(p-P_{L^2}^h p\right),\delta_K^p\nabla\left(p^h-P_{L^2}^h p\right)\right)_K\\&\le C\frac{h^{2(l+1)}}{\nu}\left\|p\right\|_{H^{l+1}(\Omega)}^2+\frac{1}{16}\sum_{K\in\mathcal{T}^h}\delta_K^p\left\|\nabla\left(p^h-P_{L^2}^h p\right)\right\|_{L^2(K)}^2.\end{aligned}$$

The term with the pressure jump vanishes for $l \ge 1$ since in this case it is $\left[\left|p-P_{L^2}^h p\right|\right]_E = 0$ for all faces by the choice of $P_{L^2}^h p$. For $l = 0$, one gets with (C.34)

$$\begin{aligned}&\sum_{E\in\mathcal{E}^h}\gamma_E\left(\left[\left|p-P_{L^2}^h p\right|\right]_E,\left[\left|p^h-P_{L^2}^h p\right|\right]_E\right)_E\\&\le C\frac{h^{2(l+1)}}{\nu}\left\|p\right\|_{H^{l+1}(\Omega)}^2+\frac{1}{16}\sum_{E\in\mathcal{E}^h}\gamma_E\left\|\left[\left|p^h-P_{L^2}^h p\right|\right]_E\right\|_{L^2(E)}^2.\end{aligned}$$

Collecting all estimates proves the statement of the theorem. ■

Remark 4.99 (Discussion of Estimate (4.131) *and Further Analytical Results)*

- It can be deduced from the error estimate (4.131) that the order of convergence for the error in the norm $\|\cdot\|_{\text{pspg}}$ is at least $\min\{k, l+1\}$. Small values of ν lead to large weights of the pressure contribution in the error bound.
- The gradient of the velocity is part of $\|\cdot\|_{\text{pspg}}$. By neglecting the other terms in $\left\|\left(\boldsymbol{u}-\boldsymbol{u}^h, p-p^h\right)\right\|_{\text{pspg}}$ and dividing by $\nu^{1/2}$, one obtains an estimate for $\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}$. The factor ν^{-1} appears in front of the pressure term in the error bound, which is the same factor as for inf-sup stable discretizations, compare (4.45). Applying Poincaré's inequality (A.12) gives a bound for the velocity error in $L^2(\Omega)$, but this bound is not optimal.
- For a family of quasi-uniform triangulations, see Definition C.9, one has $h_K \le h \le Ch_K$ for all mesh cells K and a given constant $C \ge 1$. Considering a pressure finite element space consisting only of continuous functions, using (4.130), and neglecting the velocity term in $\left\|\left(\boldsymbol{u}-\boldsymbol{u}^h, p-p^h\right)\right\|_{\text{pspg}}$ (the jump term vanishes

anyway for continuous pressure), one gets from (4.131)

$$\left\|\nabla\left(p-p^h\right)\right\|_{L^2(\Omega)} \le C\left(\nu h^{k-1}\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + h^l\|p\|_{H^{l+1}(\Omega)}\right).$$

Thus, one expects the order of convergence to be at least $\min\{k-1, l\}$ and in addition large pressure errors for large values of ν.

- Error estimates for $\|\boldsymbol{u}-\boldsymbol{u}^h\|_{L^2(\Omega)}$ and $\|p-p^h\|_{L^2(\Omega)}$ were derived in Brezzi and Douglas (1988).
- A modification of the PSPG method for continuous discrete pressure that is stable for stabilization parameters $\delta = Ch^2$ with arbitrary $C > 0$, in contrast to condition (4.125), and that allows error estimates in the standard norms of the continuous problem will be discussed in Remark 4.106.

□

Example 4.100 (Analytical Example Which Supports the Error Estimate (4.131)) Example D.3 with analytical solution is considered for the pairs P_1/P_1 and P_1/P_0 of finite element spaces. The simulations were performed on the unstructured triangular grid depicted in Fig. 4.2. Instead of presenting results for the norm $\|\cdot\|_{\mathrm{pspg}}$, as in estimate (4.131), results for the standard norms will be shown, see the comments concerning the $L^2(\Omega)$ norm of the velocity and the same norm of the pressure in Remark 4.99.

Figure 4.14 shows results for the P_1/P_1 pair of finite element spaces. The stabilization parameter in the simulations was set to be $\delta_K^p = 0.5h_K^2/\nu$, where h_K is the diameter of the mesh cell K. Small values of ν lead to large velocity errors on coarse grids. On finer grids, the errors tend to become independent of ν. One can see the expected first order of convergence for the gradient of the velocity error, which is part of the norm $\left\|\left(\boldsymbol{u}-\boldsymbol{u}^h, p-p^h\right)\right\|_{\mathrm{pspg}}$. The velocity error in $L^2(\Omega)$ is of second order convergent. Considering the pressure, large values of ν lead to large errors. One can observe an order of convergence of 1.5 in the $L^2(\Omega)$ norm. For small values of ν, the order of convergence is even higher.

Results for the P_1/P_0 pair of spaces and with $\gamma_E = 0.5h_E/\nu$ are presented in Fig. 4.15. With respect to the velocity, similar observations can be made as for the P_1/P_1 pair of spaces. However, it is not clear if the difference of the errors for different values of ν tends to vanish on finer grids. One can see this behavior for $\nu = 10^{-2}$ for the gradient and divergence, but for the error in $L^2(\Omega)$ the distance of the lines seems to be constant. The error for the pressure in $L^2(\Omega)$ converges of first order. Large values of ν lead to large errors. □

Remark 4.101 (Implementation) The PSPG term introduces a pressure-pressure coupling, it influences the velocity (ansatz) - pressure (test) coupling, and it defines a non-zero right-hand side for the pressure test functions. In detail, the matrix entries are given by

$$(B)_{ij} = b_{ij} = \sum_{K\in\mathcal{T}^h}\left[-\left(\nabla\cdot\boldsymbol{\phi}_j^h, \psi_i^h\right)_K + \left(-\nu\Delta\boldsymbol{\phi}_j^h, \delta_K^p\nabla\psi_i^h\right)_K\right],$$

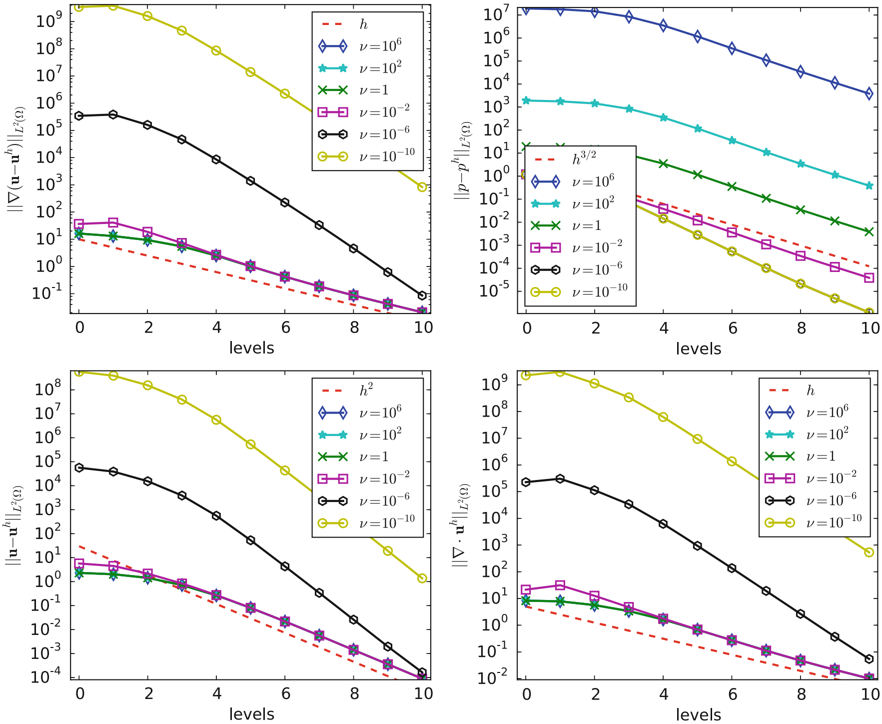

Fig. 4.14 Example 4.100. Convergence of different errors for the PSPG method with the P_1/P_1 pair of finite element spaces

$i = 1, \ldots, N_p, j = 1, \ldots, 3N_v$, and

$$(C)_{ij} = c_{ij} = \sum_{E \in \mathcal{E}^h} \gamma_E \left([|\psi_j^h|]_E, [|\psi_i^h|]_E \right)_E + \sum_{K \in \mathcal{T}^h} \left(\nabla \psi_j^h p^h, \delta_K^p \nabla \psi_i^h \right)_K,$$

$i, j = 1, \ldots, N_p$. On the right-hand side, the term

$$(\underline{f_p})_i = - \sum_{K \in \mathcal{T}^h} \left(\boldsymbol{f}, \delta_K^p \nabla \psi_i^h \right)_K, \quad i = 1, \ldots, N_p,$$

appears. The coupled system has the form

$$\begin{pmatrix} A & D \\ B & -C \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{f_p} \end{pmatrix},$$

where the matrix A is the same matrix as in the Stokes problem (4.91), it is of block-diagonal form (4.93), and the matrix D is the transposed of the matrix B from

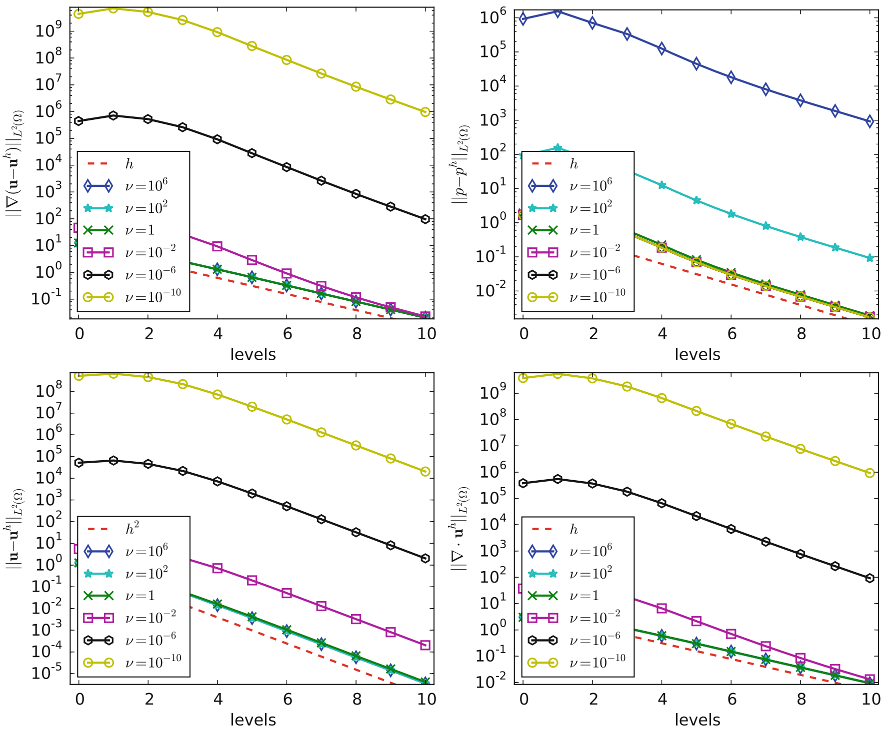

Fig. 4.15 Example 4.100. Convergence of different errors for the PSPG method with the P_1/P_0 pair of finite element spaces

the Stokes problem (4.92). In summary, the PSPG method requires the additional storage of a velocity-pressure matrix, a pressure-pressure matrix, and a right-hand side for the pressure test functions. □

Remark 4.102 (A Post-processed Weakly Divergence-free Velocity for P_1/P_0) If the PSPG method is used with the P_1/P_0 finite element, then it is possible to compute a divergence-free velocity field in $H_{\mathrm{div}}(\Omega)$ with an inexpensive post-processing step. The idea consists in adding to $\boldsymbol{u}^h$ a correction $\boldsymbol{u}^h_{\mathrm{RT}_0} \in \mathrm{RT}_0$ such that $\nabla\cdot\left(\boldsymbol{u}^h + \boldsymbol{u}^h_{\mathrm{RT}_0}\right) = 0$ in $L^2(\Omega)$. A post-processing of this type can be applied to certain stabilizations of the violation of the discrete inf-sup condition, e.g., see Barrenechea and Valentin (2010a,b, 2011). In particular, the type of stabilization which appears in the PSPG method for the P_1/P_0 pair of finite element spaces was considered in Barrenechea and Valentin (2011).

The discrete continuity equation for the $V^h/Q^h = P_1/P_0$ pair stabilized with the PSPG method (4.120), (4.121) has the form

$$\left(\nabla \cdot \boldsymbol{u}^h, q^h\right) + \sum_{E\in\mathcal{E}^h} \gamma_E \left(\left[|p^h|\right]_E, \left[|q^h|\right]_E\right)_E = 0, \tag{4.135}$$

since the Laplacian of $\boldsymbol{u}^h$ and the gradients of p^h and q^h vanish. In the two-dimensional case it is proposed in Barrenechea and Valentin (2011) to define the correction

$$\boldsymbol{u}^h_{\mathrm{RT}_0} = \sum_{E\in\mathcal{E}^h} \frac{\gamma_E}{h_E}\left(\int_E [|p^h|]_E \, ds\right)\boldsymbol{\phi}_E \tag{4.136}$$

with $\boldsymbol{u}^h_{\mathrm{RT}_0}\cdot\boldsymbol{n}_{\partial\Omega} = 0$ and

$$\boldsymbol{\phi}_E|_K = \pm\frac{h_E}{2\,|K|}\,(\boldsymbol{x}-\boldsymbol{x}_E) \in \mathrm{RT}_0(K), \tag{4.137}$$

where $\boldsymbol{x}_E$ is the node opposite to the edge E and the plus sign is used if $\boldsymbol{n}_E$ in (4.136) is the outward pointing unit normal with respect to K. By the definition of the Raviart–Thomas space of lowest order, see Example B.45, $\boldsymbol{\phi}_E\cdot\boldsymbol{n}_E$ is constant on E and $\boldsymbol{\phi}_E\cdot\boldsymbol{n}_E$ is continuous across E.

Let $\boldsymbol{n}_E$ be the outward pointing unit normal with respect to K and let $\tilde{\boldsymbol{x}}_E \in E$ (or on the straight continuation of E) be the point such that $\tilde{\boldsymbol{x}}_E - \boldsymbol{x}_E$ is perpendicular to E. Then, the area of the triangle K is given by

$$|K| = \frac{h_E\,\|\tilde{\boldsymbol{x}}_E-\boldsymbol{x}_E\|_2}{2}. \tag{4.138}$$

Since $\boldsymbol{\phi}_E\cdot\boldsymbol{n}_E$ is constant on E, one gets this constant by inserting any point on E (or on its straight continuation) in (4.137), in particular $\tilde{\boldsymbol{x}}_E$. One obtains, using (4.138), the parallelism of the vectors $\tilde{\boldsymbol{x}}_E-\boldsymbol{x}_E$ and $\boldsymbol{n}_E$, and $\|\boldsymbol{n}_E\|_2 = 1$

$$\begin{aligned}\boldsymbol{\phi}_E\cdot\boldsymbol{n}_E &= \frac{h_E}{2\,|K|}\,(\tilde{\boldsymbol{x}}_E-\boldsymbol{x}_E)\cdot\boldsymbol{n}_E = \frac{(\tilde{\boldsymbol{x}}_E-\boldsymbol{x}_E)\cdot\boldsymbol{n}_E}{\|\tilde{\boldsymbol{x}}_E-\boldsymbol{x}_E\|_2}\\ &= \frac{\|\tilde{\boldsymbol{x}}_E-\boldsymbol{x}_E\|_2\,\|\boldsymbol{n}_E\|_2}{\|\tilde{\boldsymbol{x}}_E-\boldsymbol{x}_E\|_2} = 1.\end{aligned} \tag{4.139}$$

Let K_1 and K_2 be mesh cells with the common edge E and let $\boldsymbol{n}_E$ be the outward pointing unit normal with respect to K_1. Using that the discrete pressure is piecewise constant and changing then the direction of the normal in the term coming from K_2 yields

$$\begin{aligned}&\gamma_E\left([|p^h|]_E, [|q^h|]_E\right)_E\Big|_{\boldsymbol{n}_E=\boldsymbol{n}_{\partial K_1}} \\ &= \gamma_E\, q^h|_{K_1}\int_E [|p^h|]_E\, ds\Big|_{\boldsymbol{n}_E=\boldsymbol{n}_{\partial K_1}} - \gamma_E\, q^h|_{K_2}\int_E [|p^h|]_E\, ds\Big|_{\boldsymbol{n}_E=\boldsymbol{n}_{\partial K_1}} \\ &= \gamma_E\, q^h|_{K_1}\int_E [|p^h|]_E\, ds\Big|_{\boldsymbol{n}_E=\boldsymbol{n}_{\partial K_1}} + \gamma_E\, q^h|_{K_2}\int_E [|p^h|]_E\, ds\Big|_{\boldsymbol{n}_E=\boldsymbol{n}_{\partial K_2}}.\end{aligned} \tag{4.140}$$

For an edge $\hat{E}$, it is $\boldsymbol{\phi}_E \cdot \boldsymbol{n}_{\hat{E}} = 0$ for $\hat{E} \neq E$ since either $\boldsymbol{\phi}_E$ vanishes on $\hat{E}$ or it is perpendicular to $\boldsymbol{n}_{\hat{E}}$, because $\{\boldsymbol{\phi}_E\}$ forms a local basis and $\boldsymbol{\phi}_E \cdot \boldsymbol{n}_{\hat{E}}$ are the corresponding functionals. Hence, one obtains with (4.136) and (4.139)

$$\boldsymbol{u}^h_{\mathrm{RT}_0} \cdot \boldsymbol{n}_{\hat{E}}\big|_{\hat{E}} = \sum_{E\in\mathcal{E}^h} \frac{\gamma_E}{h_E}\left(\int_E [|p^h|]_E \, ds\right) \boldsymbol{\phi}_E\Bigg|_{\hat{E}} \cdot \boldsymbol{n}_{\hat{E}} = \frac{\gamma_{\hat{E}}}{h_{\hat{E}}} \int_{\hat{E}} [|p^h|]_{\hat{E}} \, ds. \quad (4.141)$$

Considering the contribution of K_1 in (4.140), inserting (4.141) in (4.140), and using that q^h and $\boldsymbol{u}^h_{\mathrm{RT}_0} \cdot \boldsymbol{n}_E$ are both constant leads to

$$\gamma_E \left([|p^h|]_E \, , q^h|_{K_1}\right)_E = q^h|_{K_1} h_E \left(\boldsymbol{u}^h_{\mathrm{RT}_0} \cdot \boldsymbol{n}_E\right) = \int_E q^h|_{K_1} \left(\boldsymbol{u}^h_{\mathrm{RT}_0} \cdot \boldsymbol{n}_E\right) \, ds. \quad (4.142)$$

Writing now (4.135) as a sum over the mesh cells, using (4.140), (4.142), applying integration by parts, and using that q^h is piecewise constant and that $\boldsymbol{u}^h + \boldsymbol{u}^h_{\mathrm{RT}_0} \in L^2(\Omega)$ gives for all $q^h \in Q^h$

$$\begin{aligned} 0 &= \sum_{K\in\mathcal{T}^h}\left[\left(\nabla\cdot\boldsymbol{u}^h, q^h\right)_K + \sum_{E\in\partial K} \gamma_E \left([|p^h|]_E \, , q^h|_K\right)_E\right] \\ &= \sum_{K\in\mathcal{T}^h}\left[\left(\nabla\cdot\boldsymbol{u}^h, q^h\right)_K + \int_{\partial K} q^h|_K \left(\boldsymbol{u}^h_{\mathrm{RT}_0}\cdot\boldsymbol{n}_{\partial K}\right) \, ds\right] \\ &= \sum_{K\in\mathcal{T}^h}\left[\left(\nabla\cdot\boldsymbol{u}^h, q^h\right)_K + \left(\nabla\cdot\boldsymbol{u}^h_{\mathrm{RT}_0}, q^h\right)_K\right] \\ &= \sum_{K\in\mathcal{T}^h}\left(\nabla\cdot\left(\boldsymbol{u}^h+\boldsymbol{u}^h_{\mathrm{RT}_0}\right), q^h\right)_K = \left(\nabla\cdot\left(\boldsymbol{u}^h+\boldsymbol{u}^h_{\mathrm{RT}_0}\right), q^h\right). \end{aligned}$$

Hence, $\nabla \cdot \left(\boldsymbol{u}^h + \boldsymbol{u}^h_{\mathrm{RT}_0}\right)$ is perpendicular to Q^h. On the other hand, $\nabla \cdot \left(\boldsymbol{u}^h + \boldsymbol{u}^h_{\mathrm{RT}_0}\right)$ is a piecewise constant function and integration by parts and using the continuity of the normal components of $\boldsymbol{u}^h + \boldsymbol{u}^h_{\mathrm{RT}_0}$ across edges shows $\left(\nabla \cdot \left(\boldsymbol{u}^h + \boldsymbol{u}^h_{\mathrm{RT}_0}\right), 1\right) = 0$, such that $\nabla \cdot \left(\boldsymbol{u}^h + \boldsymbol{u}^h_{\mathrm{RT}_0}\right) \in Q^h$. Consequently, it follows that $\nabla \cdot \left(\boldsymbol{u}^h + \boldsymbol{u}^h_{\mathrm{RT}_0}\right) = 0$ in Ω and since by construction $\left(\boldsymbol{u}^h + \boldsymbol{u}^h_{\mathrm{RT}_0}\right) \in H(\mathrm{div}, \Omega)$, one gets that $\left(\boldsymbol{u}^h + \boldsymbol{u}^h_{\mathrm{RT}_0}\right) \in H_{\mathrm{div}}(\Omega)$. □

Example 4.103 (P_1/P_0 with Post-processing) This example continues Example 4.100. Figure 4.16 presents results for the P_1/P_0 pair of spaces with the post-processing described in Remark 4.102. Comparing the velocity errors of $\boldsymbol{u}^h + \boldsymbol{u}^h_{\mathrm{RT}_0}$ in $L^2(\Omega)$ with those of $\boldsymbol{u}^h$ presented in Fig. 4.15, one can see that there are almost no differences. However, the divergence of $\boldsymbol{u}^h + \boldsymbol{u}^h_{\mathrm{RT}_0}$ is almost zero, at least for large values of ν. Even for $\nu \in \{10^{-6}, 10^{-10}\}$, the divergence is smaller by more than ten orders of magnitude compared with the divergence of $\boldsymbol{u}^h$. The larger

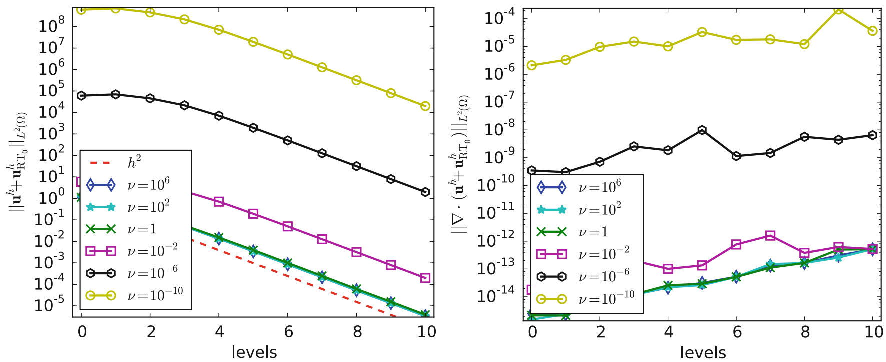

Fig. 4.16 Example 4.103. Convergence of errors for the post-processed velocity solution computed with the PSPG method and the P_1/P_0 pair of finite element spaces

values for $\nu \in \{10^{-6}, 10^{-10}\}$ are due to round-off errors coming from dealing with the large values of the divergence of $\boldsymbol{u}^h$ itself. □

Remark 4.104 (Part of Residual-Based Stabilization for the Oseen Equations) The PSPG method will be a part of a residual-based stabilization for the Oseen equations, see Sect. 5.3.2. □

4.5.2 Some Other Stabilized Methods

Remark 4.105 (Framework for Stabilizations Using the Residual) In Bochev and Gunzburger (2004), framework is presented for stabilizations with respect to the discrete inf-sup condition that use the residual. The PSPG method from Sect. 4.5.1, its modification from Remark 4.106, the Galerkin least squares method from Remark 4.107, and the method from Douglas and Wang (1989) from Remark 4.108 fit into this framework. □

Remark 4.106 (An Absolutely Stable Modification of the PSPG Method) In Bochev and Gunzburger (2004), a modification of the PSPG method for continuous pressure approximations is proposed that allows to choose the stabilization parameter as $\delta = Ch^2$ with any $C > 0$. This method is called absolutely stable. Recall that for the PSPG method studied in Sect. 4.5.1 there is the upper bound (4.125). The modification consists in using instead of

$$\sum_{K\in\mathcal{T}^h} \left(-\nu\Delta\boldsymbol{u}^h, \delta_K^p \nabla q^h\right)_K \qquad (4.143)$$

as in (4.120), a stabilization term of the form

$$\sum_{K\in\mathcal{T}^h} \left(-\nu\Delta^h \boldsymbol{u}^h, \delta_K^p \nabla q^h\right)_K, \tag{4.144}$$

where the discrete Laplacian $\boldsymbol{z}^h = \Delta^h \boldsymbol{u}^h \in V^h$ is the solution of

$$\left(\boldsymbol{z}^h, \boldsymbol{v}^h\right) = \left(\nabla \boldsymbol{u}^h, \nabla \boldsymbol{v}^h\right) \quad \forall\ \boldsymbol{v}^h \in V^h. \tag{4.145}$$

Thus, the modified PSPG method requires the additional solution of problem (4.145), which is a linear system with the mass matrix as system matrix.

The numerical analysis in Bochev and Gunzburger (2004) shows the mesh-independent absolute stability of the modified PSPG method and its optimal convergence with respect to $\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}$ and $\left\|p-p^h\right\|_{L^2(\Omega)}$. In addition, it is emphasized that in the case of piecewise linear finite elements, expression (4.143) vanishes whereas expression (4.144) does not vanish, which might lead to a higher accuracy of the modified method. □

Remark 4.107 (The Galerkin Least Squares (GLS) Method) The Galerkin Least Squares (GLS) method uses, like the PSPG method (4.119)–(4.121), the residual of the strong form of the equation. In contrast to the PSPG method, the operator of the strong form of the equation appears also for the test functions. Hence, the application of a GLS method is a little bit more expensive than the use of the PSPG method.

The GLS method was proposed in Hughes and Franca (1987). It has the form: Find $\left(\boldsymbol{u}^h, p^h\right) \in V^h \times Q^h$ such that

$$A_{\text{gls}}\left(\left(\boldsymbol{u}^h, p^h\right), \left(\boldsymbol{v}^h, q^h\right)\right) = L_{\text{gls}}\left(\left(\boldsymbol{v}^h, q^h\right)\right) \quad \forall\ \left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h, \tag{4.146}$$

with

$$\begin{aligned}
&A_{\text{gls}}\left((\boldsymbol{u}, p), (\boldsymbol{v}, q)\right) \\
&\quad = \nu\left(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}\right) - \left(\nabla\cdot\boldsymbol{v}, p\right) - \left(\nabla\cdot\boldsymbol{u}, q\right) \\
&\qquad - \sum_{E\in\mathcal{E}^h} \gamma_E\left([|p|]_E, [|q|]_E\right)_E - \sum_{K\in\mathcal{T}^h}\left(-\nu\Delta\boldsymbol{u} + \nabla p, \delta_K^p\left(-\nu\Delta\boldsymbol{v} + \nabla q\right)\right)_K, \\
&L_{\text{gls}}\left((\boldsymbol{v}, q)\right) = \left(\boldsymbol{f}, \boldsymbol{v}\right) - \sum_{K\in\mathcal{T}^h}\left(\boldsymbol{f}, \delta_K^p\left(-\nu\Delta\boldsymbol{v} + \nabla q\right)\right)_K.
\end{aligned} \tag{4.147}$$

This method is symmetric. Optimal error estimates can be proved for $\delta_K^p = \mathcal{O}\left(h_K^2/\nu\right)$ and $\gamma_E = \mathcal{O}\left(h_E/\nu\right)$ (in the case of a discontinuous pressure approximation), see Hughes and Franca (1987) for details. The stability estimate for the solution requires an upper bound of the stabilization parameter, hence this method is not absolutely stable. □

Remark 4.108 (An Absolutely Stable Method from Douglas and Wang (1989)) A method that looks similar to the GLS method (4.146), (4.147) was proposed in Douglas and Wang (1989): Find $\left(\boldsymbol{u}^h, p^h\right) \in V^h \times Q^h$ such that

$$A_{\mathrm{DW}}\left(\left(\boldsymbol{u}^h, p^h\right), \left(\boldsymbol{v}^h, q^h\right)\right) = L_{\mathrm{DW}}\left(\left(\boldsymbol{v}^h, q^h\right)\right) \quad \forall \ \left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h, \tag{4.148}$$

with

$$\begin{aligned}
&A_{\mathrm{DW}}\left((\boldsymbol{u}, p), (\boldsymbol{v}, q)\right) \\
&= \nu\left(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}\right) - \left(\nabla \cdot \boldsymbol{v}, p\right) + \left(\nabla \cdot \boldsymbol{u}, q\right) \\
&\quad + \sum_{E \in \mathcal{E}^h} \gamma_E \left([|p|]_E, [|q|]_E\right)_E + \sum_{K \in \mathcal{T}^h} \left(-\nu \Delta \boldsymbol{u} + \nabla p, \delta_K^p \left(-\nu \Delta \boldsymbol{v} + \nabla q\right)\right)_K, \\
&L_{\mathrm{DW}}\left((\boldsymbol{v}, q)\right) = (\boldsymbol{f}, \boldsymbol{v}) + \sum_{K \in \mathcal{T}^h} \left(\boldsymbol{f}, \delta_K^p \left(-\nu \Delta \boldsymbol{v} + \nabla q\right)\right)_K.
\end{aligned} \tag{4.149}$$

The difference of (4.146), (4.147) and (4.148), (4.149) is just the sign that is used to incorporate the discrete continuity equation in the discrete momentum equation. Method (4.148), (4.149) corresponds to the form (4.3).

Despite the methods (4.146), (4.147) and (4.148), (4.149) look similar, their properties differ substantially, e.g., see Franca et al. (1993), Barth et al. (2004). Method (4.148), (4.149) is non-symmetric and it is absolutely stable. Optimal error estimates for the parameter choices $\delta_K^p = \mathcal{O}\left(h_K^2/\nu\right)$ and $\gamma_E = \mathcal{O}\left(h_E/\nu\right)$ (in the case of a discontinuous discrete pressure) were derived in Douglas and Wang (1989). □

Remark 4.109 (A Framework for Stabilizations Using only the Pressure) In Brezzi and Fortin (2001), a framework for constructing stabilizations with respect to the discrete inf-sup condition (3.51) was derived. In this framework, the stabilization term was chosen such that there is a minimal perturbation with respect to the original problem. Concrete realizations given in Brezzi and Fortin (2001) consider stabilization terms that only contain the pressure and not, e.g., the residual. several stabilized methods presented in this section fit into this framework, e.g., the methods from Brezzi and Pitkäranta (1984) and Codina and Blasco (1997). □

Remark 4.110 (Stabilization from Brezzi and Pitkäranta (1984)) The first stabilization method for circumventing the discrete inf-sup condition (3.51) was proposed in Brezzi and Pitkäranta (1984). This proposal was for the P_1/P_1 pair of finite element spaces and it has the form: Find $(\boldsymbol{u}^h, p^h) \in V^h \times Q^h = P_1 \times P_1$ such that

$$\begin{aligned}
\nu\left(\nabla \boldsymbol{u}^h, \nabla \boldsymbol{v}^h\right) - \left(\nabla \cdot \boldsymbol{v}^h, p^h\right) &= \left(\boldsymbol{f}, \boldsymbol{v}^h\right) \ \forall \ \boldsymbol{v}^h \in V^h, \\
-\left(\nabla \cdot \boldsymbol{u}^h, q^h\right) - \sum_{K \in \mathcal{T}^h} \left(\nabla p^h, \delta_K^p \nabla q^h\right)_K &= 0 \qquad \forall \ q^h \in Q^h.
\end{aligned} \tag{4.150}$$

In the case of a uniform family of triangulations, optimal order convergence of (4.150) with respect to $\left\|\nabla\left(\boldsymbol{u} - \boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}$ and $\left\|p - p^h\right\|_{L^2(\Omega)}$ was proved for

$\delta_K^p = \mathcal{O}\left(h^2\right)$ (for $\nu = 1$). Method (4.150) fits into the framework discussed in Remark 4.109.

Consider the case $\delta_K^p = \delta$ and applying integration by parts of the continuity equation of (4.150) gives

$$\begin{aligned}
&-\left(\nabla\cdot\boldsymbol{u}^h, q^h\right) - \sum_{K\in\mathcal{T}^h}\left(\nabla p^h, \delta_K^p \nabla q^h\right)_K \\
&= \left(-\nabla\cdot\boldsymbol{u}^h + \delta\Delta p^h, q^h\right) - \delta\int_{\partial\Omega}\left(\nabla p^h\cdot\boldsymbol{n}\right) q^h\, d\boldsymbol{s}.
\end{aligned}$$

With the same arguments as in Remark 4.93, one finds that the artificial boundary condition

$$\delta\left(\nabla p^h\cdot\boldsymbol{n}\right) = 0 \quad \text{on } \partial\Omega$$

for the discrete pressure is introduced with this method. □

Remark 4.111 (Stabilization with Fluctuations of the Pressure Gradient) It was shown in Codina and Blasco (1997) that it is not necessary t use the full gradient of the discrete pressure, as in (4.150), for constructing a stable method. Let $\overline{V^h}$ be the velocity finite element space with the same polynomials as V^h but without prescribed boundary conditions. Then, the method proposed in Codina and Blasco (1997) has the form: Find $\left(\boldsymbol{u}^h, p^h, \overline{\nabla p^h}\right) \in V^h\times Q^h\times\overline{V^h}$ such that

$$\begin{aligned}
\nu\left(\nabla\boldsymbol{u}^h, \nabla\boldsymbol{v}^h\right) - \left(\nabla\cdot\boldsymbol{v}^h, p^h\right) &= \left(\boldsymbol{f}, \boldsymbol{v}^h\right) & \forall\ \boldsymbol{v}^h \in V^h,\\
-\left(\nabla\cdot\boldsymbol{u}^h, q^h\right) - \sum_{K\in\mathcal{T}^h}\left(\nabla p^h - \overline{\nabla p^h}, \delta_K^p\nabla q^h\right)_K &= 0 & \forall\ q^h \in Q^h,\\
\left(\nabla p^h - \overline{\nabla p^h}, \overline{\boldsymbol{v}^h}\right) &= 0 & \forall\ \overline{\boldsymbol{v}^h} \in \overline{V^h}.
\end{aligned} \tag{4.151}$$

The last equation of (4.151) states that $\overline{\nabla p^h}$ is the $L^2(\Omega)$ projection of ∇p^h onto $\overline{V^h}$. Thus, $\overline{\nabla p^h}$ can be interpreted to represent large scales of ∇p^h. The difference $(\nabla p^h - \overline{\nabla p^h}$ is called fluctuations and these fluctuations appear in the stability term in the discrete continuity equation. This method fits into the framework presented in Remark 4.109.

In the case of a family of quasi-uniform triangulations, $\delta_K^p = \delta$ and for the $\delta = \mathcal{O}\left(h^2\right)$, the stability of the finite element solution, a discrete inf-sup condition of form (3.128), and optimal error estimates for $\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}$, $\left\|\nabla\left(p-p^h\right)\right\|_{L^2(\Omega)}$, and $\left\|\nabla p - \overline{\nabla p^h}\right\|_{L^2(\Omega)}$ were proved in Codina and Blasco (1997). Extensions of the analysis that allow locally chosen stabilization parameters and to the steady-state Navier–Stokes equations can be found in Codina and Blasco (2000). □

Remark 4.112 (A Local Projection Stabilization (LPS) Method) The computation of the global $L^2(\Omega)$ projection in method (4.151) introduces some computational overhead. In Becker and Braack (2001), it was proposed to replace the global projection by local projections, whose computation is less expensive, leading to the so-called Local Projection Stabilization (LPS) method. Later, it was observed that the local projection approach can be used for stabilizing further potential instabilities that might occur in the simulation of incompressible flows. The LPS method became popular to be applied to the Oseen equations and the Navier–Stokes equations. This method will be discussed in some detail in the context of the Oseen equations, starting from Remark 5.52. □

Remark 4.113 (A Stabilization Using Fluctuations of the Pressure) A stabilization with respect to the discrete inf-sup condition was proposed in Dohrmann and Bochev (2004) that uses fluctuations of the pressure itself, instead of the gradient of the pressure as the methods discussed in Remarks 4.111 and 4.112. Let $Q^h = P_k$ or $Q^h = Q_k$ with $k \geq 1$. The definition of the method uses the $L^2(\Omega)$ projection $P_{L^2}^{k-1} : Q^h \to P_{k-1}^{\text{disc}}$ onto the discontinuous piecewise polynomial space of degree $k-1$. Because the projection space is discontinuous, the action of $P_{L^2}^{k-1}$ can be computed locally, i.e., mesh cell by mesh cell. Then, the continuity equation of the stabilized method proposed in Dohrmann and Bochev (2004) has the form

$$-\left(\nabla \cdot \boldsymbol{u}^h, q^h\right) - \frac{1}{\nu}\left(p^h - P_{L^2}^{k-1} p^h, q^h - P_{L^2}^{k-1} q^h\right) = 0 \quad \forall\, q^h \in Q^h. \tag{4.152}$$

Note that this method is parameter-free.

A numerical analysis of this method for the lowest order pairs P_1/P_1 and Q_1/Q_1 is presented in Bochev et al. (2006), where in contrast to (4.152) the inverse of the viscosity does not appear in the stabilization term. In addition, an extension of the method to the pairs P_1/P_0 and Q_1/Q_0 is introduced. It is shown that in all cases the method is unconditionally stable and optimal error bounds, i.e., linear convergence, for $\left\|\nabla\left(\boldsymbol{u} - \boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}$ and $\left\|p - p^h\right\|_{L^2(\Omega)}$ are proved. □

4.6 Improving the Conservation of Mass, Divergence-Free Finite Element Solutions

Remark 4.114 (Motivation) The finite element velocity field is only discretely divergence-free. The violation of the divergence-free constraint, i.e., of the conservation of mass, might be quite large, e.g., see Example 4.31, and the convergence is usually only of the same order as for the error of the gradient, see Fig. 4.4. However, there are strong reasons why one is interested in computing weakly divergence-free finite element velocity solutions or at least to reduce the violation of the divergence constraint.

- A violation of the conservation of mass is not acceptable in many applications.
- The property of the finite element velocity to be not weakly divergence-free leads to the fact that the velocity errors depend on the pressure, compare (4.35)

and (4.38) for discretely divergence-free velocities with (4.51) and (4.53) for weakly divergence-free velocities.

A post-processing technique for computing a weakly divergence-free solution for a low order discretization that does not satisfy the discrete inf-sup condition was already presented in Remark 4.102.

This section presents techniques for reducing the violation of the conservation of mass or even for computing weakly divergence-free solutions in the case of inf-sup stable pairs of finite element spaces. More details can be found in the review paper (John et al. 2015). □

4.6.1 The Grad-Div Stabilization

Remark 4.115 (Literature) The grad-div stabilization introduces a penalty term in the Galerkin finite element formulation which penalizes the violation of mass conservation. Its inclusion in the analysis of incompressible flow problems can be found already in Franca and Hughes (1988), Hansbo and Szepessy (1990). A finite element error analysis was performed often in the context of the Oseen equations (2.32), e.g., see Tobiska and Verfürth (1996), Gelhard et al. (2005), Roos et al. (2008), Olshanskii et al. (2009). The case of the Stokes equations was considered in Olshanskii and Reusken (2004), Jenkins et al. (2014). □

Remark 4.116 (The Grad-Div Stabilization) The penalty or stabilization term introduces a set of user-chosen parameters. For obtaining an optimal order of convergence, the stabilization parameters have to be chosen appropriately. To highlight the dependency of these parameters not only on the mesh size but also on the data of the problem, the scaled Stokes equations (4.44) will be studied. The weak form of these equations is: Find $(\boldsymbol{u}, p) \in V \times Q$ such that

$$\begin{aligned} \nu(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, p) &= \langle \boldsymbol{f}, \boldsymbol{v}\rangle_{V',V} \ \ \forall\ \boldsymbol{v} \in V, \\ -(\nabla \cdot \boldsymbol{u}, q) &= 0 \qquad\qquad\quad \forall\ q \in Q. \end{aligned} \tag{4.153}$$

In this section, only the case of conforming finite element spaces will be considered. Then, the grad-div stabilization method reads as follows: Find $\left(\boldsymbol{u}^h, p^h\right) \in V^h \times Q^h$ such that

$$\begin{aligned} \nu \left(\nabla \boldsymbol{u}^h, \nabla \boldsymbol{v}^h\right) &- \left(\nabla \cdot \boldsymbol{v}^h, p^h\right) \\ + \sum_{K\in\mathcal{T}^h} \mu_K \left(\nabla \cdot \boldsymbol{u}^h, \nabla \cdot \boldsymbol{v}^h\right)_K &= \langle \boldsymbol{f}, \boldsymbol{v}^h\rangle_{V',V} \ \ \forall\ \boldsymbol{v}^h \in V^h, \\ - \left(\nabla \cdot \boldsymbol{u}^h, q^h\right) &= 0 \qquad\qquad\quad \forall\ q^h \in Q^h, \end{aligned} \tag{4.154}$$

where $\{\mu_K\}$ with $\mu_K \geq 0$ are the stabilization parameters.

Let

$$a^h\left(\boldsymbol{u}^h,\boldsymbol{v}^h\right) = \nu\left(\nabla\boldsymbol{u}^h,\nabla\boldsymbol{v}^h\right) + \sum_{K\in\mathcal{T}^h}\mu_K\left(\nabla\cdot\boldsymbol{u}^h,\nabla\cdot\boldsymbol{v}^h\right)_K \quad \forall\,\boldsymbol{u}^h,\boldsymbol{v}^h\in V^h.$$

The first term of this bilinear form is symmetric and positive definite and the second term is symmetric and positive semi-definite. Hence, $a^h(\cdot,\cdot)$ is V^h-elliptic and in particular V^h_{div}-elliptic. The existence and uniqueness of a solution for pairs of finite element spaces that satisfy the discrete inf-sup condition (3.51) follows analogously to the proof of Theorem 4.6.

Define

$$\mu = \max_{K\in\mathcal{T}^h}\mu_K.$$

□

Remark 4.117 (On the Grad-Div Term)

- Applying integration by parts and noting that the boundary integral vanishes since $\nabla\cdot\boldsymbol{u}=0$ on the boundary, one obtains for $\mu_K=\mu=\mathrm{const}$

$$\boldsymbol{0} = \mu\left(\nabla\cdot\boldsymbol{u},\nabla\cdot\boldsymbol{v}\right) = -\mu\left(\nabla\left(\nabla\cdot\boldsymbol{u}\right),\boldsymbol{v}\right) \quad \forall\,\boldsymbol{u},\boldsymbol{v}\in V.$$

  The notation grad-div stabilization term comes from the term on the right-hand side of this equation.
- A grad-div term appears already in the derivation of the Navier–Stokes equations, see (2.20), which reads for constant viscosity

$$\nabla\cdot\left(\left(\zeta-\frac{2\mu}{3}\right)\nabla\cdot\boldsymbol{v}\mathbb{I}\right) = \left(\zeta-\frac{2\mu}{3}\right)\nabla\left(\nabla\cdot\boldsymbol{v}\right),$$

  where μ is here the dynamic viscosity. In the process of deriving the non-dimensional equations, this term becomes

$$\frac{L}{U^2}\left(\frac{\zeta}{\rho}-\frac{2\nu}{3}\right)\nabla\left(\nabla\cdot\boldsymbol{u}\right).$$

  The parameter in front of this grad-div term is just a constant.
- A scaled grad-div stabilization term occurs also in the so-called augmented Lagrangian-based preconditioner for solving linear saddle point problems, see Remark 9.35.

□

Lemma 4.118 (Vanishing of $\left\|\nabla\cdot\boldsymbol{u}^h\right\|_{L^2(\Omega)}$ for the Stabilization Parameters Tending to Infinity) *Consider a fixed triangulation $\mathcal{T}^h$ and set $\mu_{\min} = \min_{K\in\mathcal{T}^h}\mu_K$. Then it holds*

$$\lim_{\mu_{\min}\to\infty}\left\|\nabla\cdot\boldsymbol{u}^h\right\|_{L^2(\Omega)} = 0.$$

The convergence with respect to $\mu_{\min}$ *is of order* 0.5 *with a constant independent of the mesh width and of order* 1 *with a constant that might depend on the mesh width.*

Proof To avoid some technical details, the proof is presented for $\mu_{\min} = \mu = \mu_K$ for all $K \in \mathcal{T}^h$.

Using the solution $\boldsymbol{u}^h$ as test function in (4.154), applying the estimate for the dual pairing, and utilizing Young's inequality (A.5) gives

$$\begin{aligned}&\nu \left\|\nabla \boldsymbol{u}^h\right\|_{L^2(\Omega)}^2 + \mu \left\|\nabla \cdot \boldsymbol{u}^h\right\|_{L^2(\Omega)}^2 \\ &= \langle \boldsymbol{f}, \boldsymbol{u}^h\rangle_{V',V} \le \|\boldsymbol{f}\|_{H^{-1}(\Omega)} \left\|\nabla \boldsymbol{u}^h\right\|_{L^2(\Omega)} \le \frac{1}{4\nu} \|\boldsymbol{f}\|_{H^{-1}(\Omega)}^2 + \nu \left\|\nabla \boldsymbol{u}^h\right\|_{L^2(\Omega)}^2 .\end{aligned}$$

Consequently, it is

$$\left\|\nabla \cdot \boldsymbol{u}^h\right\|_{L^2(\Omega)} \le \frac{1}{2\nu^{1/2}\mu^{1/2}} \|\boldsymbol{f}\|_{H^{-1}(\Omega)},$$

which yields the first statement of the lemma.

The proof of the second statement was performed in Galvin et al. (2012). It starts with defining the space

$$V^h_{\text{div,div}} = V_{\text{div}} \cap V^h_{\text{div}} \tag{4.155}$$

and its orthogonal complement in V^h_{div}

$$\left(V^h_{\text{div,div}}\right)^{\perp} = \left\{\boldsymbol{v}^h \in V^h_{\text{div}} \ : \ \left(\nabla \boldsymbol{v}^h, \nabla \boldsymbol{w}^h\right) = 0 \ \forall\ \boldsymbol{w}^h \in V^h_{\text{div,div}}\right\}.$$

In this way, one obtains a decomposition

$$V^h_{\text{div}} = V^h_{\text{div,div}} \oplus \left(V^h_{\text{div,div}}\right)^{\perp}.$$

By construction, all non-trivial weakly divergence-free functions from V^h_{div} belong to $V^h_{\text{div,div}}$. Since the only weakly divergence-free function in $\left(V^h_{\text{div,div}}\right)^{\perp}$ is $\boldsymbol{v}^h = \boldsymbol{0}$, it follows that $\left\|\nabla \cdot \boldsymbol{v}^h\right\|_{L^2(\Omega)}$ defines a norm in $\left(V^h_{\text{div,div}}\right)^{\perp}$. Clearly, $\left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)}$ is also a norm in $\left(V^h_{\text{div,div}}\right)^{\perp}$. Because all norms are equivalent in finite-dimensional spaces, see Remark A.8, there is a constant $C(h)$, which might depend on the velocity finite element space and with this on the mesh width h, such that

$$\left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)} \le C(h) \left\|\nabla \cdot \boldsymbol{v}^h\right\|_{L^2(\Omega)} \quad \forall\ \boldsymbol{v}^h \in \left(V^h_{\text{div,div}}\right)^{\perp}. \tag{4.156}$$

Now, the solution $\boldsymbol{u}^h \in V^h_{\text{div}}$ of (4.154) is decomposed into

$$\boldsymbol{u}^h = \boldsymbol{u}^h_{\text{div,div}} + \boldsymbol{u}^h_{\text{div,div},\perp}, \quad \boldsymbol{u}^h_{\text{div,div}} \in V^h_{\text{div,div}}, \boldsymbol{u}^h_{\text{div,div},\perp} \in \left(V^h_{\text{div,div}}\right)^{\perp}.$$

Inserting this decomposition in (4.154), using as test function $\boldsymbol{v}^h = \boldsymbol{u}^h_{\mathrm{div,div},\perp}$, observing that $\nabla \cdot \boldsymbol{u}^h_{\mathrm{div,div}} = 0$ almost everywhere and that $\left(\nabla \boldsymbol{u}^h_{\mathrm{div,div}}, \nabla \boldsymbol{u}^h_{\mathrm{div,div},\perp}\right) = 0$ by the definition of $\left(V^h_{\mathrm{div,div}}\right)^\perp$, applying the estimate of the dual pairing, and using (4.156), one obtains

$$\begin{aligned}&\nu \left\|\nabla \boldsymbol{u}^h_{\mathrm{div,div},\perp}\right\|^2_{L^2(\Omega)} + \mu \left\|\nabla \cdot \boldsymbol{u}^h_{\mathrm{div,div},\perp}\right\|^2_{L^2(\Omega)} \\ &\le \|\boldsymbol{f}\|_{H^{-1}(\Omega)} \left\|\nabla \boldsymbol{u}^h_{\mathrm{div,div},\perp}\right\|_{L^2(\Omega)} \le C(h) \|\boldsymbol{f}\|_{H^{-1}(\Omega)} \left\|\nabla \cdot \boldsymbol{u}^h_{\mathrm{div,div},\perp}\right\|_{L^2(\Omega)} .\end{aligned}$$

From this estimate, it follows that

$$\left\|\nabla \cdot \boldsymbol{u}^h\right\|_{L^2(\Omega)} = \left\|\nabla \cdot \boldsymbol{u}^h_{\mathrm{div,div},\perp}\right\|_{L^2(\Omega)} \le \frac{C(h)}{\mu} \|\boldsymbol{f}\|_{H^{-1}(\Omega)} ,$$

which is the second statement of the lemma. ■

Remark 4.119 (On the Choice of the Stabilization Parameters)

- A detailed discussion of the behavior of $C(h)$ can be found in Galvin et al. (2012).
- An interesting theoretical result from Case et al. (2011) is that for $\mu \to \infty$ the Taylor–Hood pair P_2/P_1 of finite element spaces becomes equivalent to the Scott–Vogelius pair P_2/P_1^{disc}.
- However, the situation $\mu \to \infty$ is generally not relevant in practice. There, a good choice of the stabilization parameters should be governed by obtaining optimal error estimates with respect to some norm, e.g., see Remark 4.125 below. In addition, using very large stabilization parameters might lead to other difficulties like a large condition number of the arising matrix.

□

Definition 4.120 (Optimal Approximation Property of a Sequence of Divergence-free Subspaces) Consider a quasi-uniform family of triangulations $\{\mathcal{T}^h\}_{h>0}$ with characteristic mesh size h. If for all $\boldsymbol{v} \in V_{\mathrm{div}} \cap H^{k+1}(\Omega)$ there exists a sequence $\left\{\boldsymbol{v}^h\right\} \in V^h_{\mathrm{div,div}}$, where $V^h_{\mathrm{div,div}}$ is defined in (4.155), with

$$\left\|\nabla\left(\boldsymbol{v} - \boldsymbol{v}^h\right)\right\|_{L^2(\Omega)} \le C_{\mathrm{div}} h^k \|\boldsymbol{v}\|_{H^{k+1}(\Omega)} \tag{4.157}$$

and C_{div} independent of h, then the sequence of spaces $\left\{V^h_{\mathrm{div,div}}\right\}$ is said to possess optimal approximation property with respect to the space V_{div}. □

Remark 4.121 (Existence of a Sequence of Divergence-free Subspaces with Optimal Approximation Property) Whether or not there exists a sequence of divergence-free subspaces with optimal approximation property depends on the pair of inf-sup stable finite element spaces and on the underlying triangulation of the domain. Several

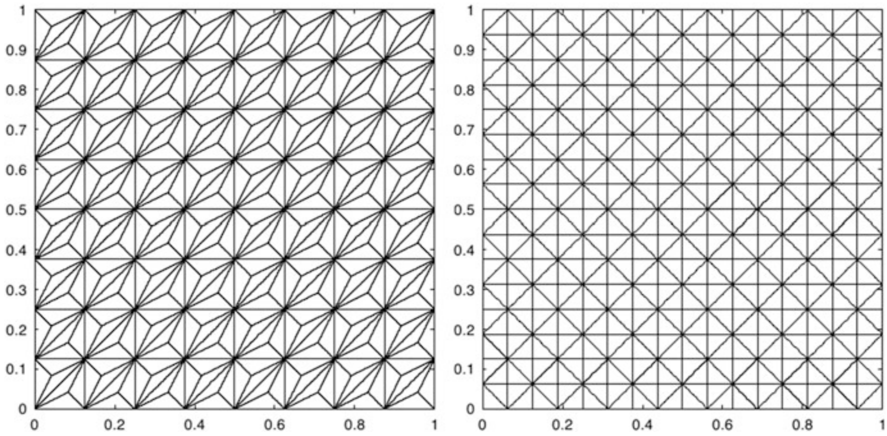

Fig. 4.17 Barycentric-refined simplicial grid (*left*) and Union Jack grid (*right*) on the unit square

combinations of pairs and triangulations are known to possess such a sequence of divergence-free subspaces. Some examples are as follows:

- Taylor–Hood pair of spaces P_k/P_{k-1} with $k \geq d$ on barycentric-refined simplicial grids, Arnold and Qin (1992), Qin (1994), Zhang (2005). The starting point for the construction of such grids is a standard grid consisting of simplices. Then, each simplex is divided into smaller simplices by connecting its barycenter with its vertices, see Fig. 4.17 for a two-dimensional example.
- MINI element on Union Jack grids (criss-cross grids), Zhang (2009). Union Jack grids are two-dimensional grids that are build from four families of parallel lines, see Fig. 4.17.

Altogether, in special cases there exists a sequence of divergence-free subspaces with optimal approximation property. In the general situation, one cannot expect such a sequence to exist. □

Theorem 4.122 (Finite Element Error Estimate for the $L^2(\Omega)$ Norm of the Gradient of the Velocity) *Let the assumptions of Theorem 4.21 be satisfied, let* $(\boldsymbol{u}, p)$ *be the solution of* (4.153) *and let* $\left(\boldsymbol{u}^h, p^h\right)$ *be the solution of* (4.154)*. Then, the error in the* $L^2(\Omega)$ *norm of the gradient of the velocity is bounded by*

$$\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}^2 \leq \inf_{\boldsymbol{v}^h\in V_{\mathrm{div}}^h}\left(4\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{v}^h\right)\right\|_{L^2(\Omega)}^2+2\frac{\mu}{\nu}\left\|\nabla\cdot\boldsymbol{v}^h\right\|_{L^2(\Omega)}^2\right) + \frac{2}{\mu\nu}\inf_{q^h\in Q^h}\left\|p-q^h\right\|_{L^2(\Omega)}^2. \tag{4.158}$$

Proof The proof is similar to the proof of Theorem 4.21. It starts with the error decomposition $\boldsymbol{u}-\boldsymbol{u}^h = \left(\boldsymbol{u}-\boldsymbol{v}^h\right) - \left(\boldsymbol{u}^h-\boldsymbol{v}^h\right) = \boldsymbol{\eta}-\boldsymbol{\phi}^h$, where $\boldsymbol{v}^h \in V_{\mathrm{div}}^h$ is arbitrary. First, by the triangle inequality and Young's inequality (A.5), one obtains

$$\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}^2 \leq 2\left\|\nabla\boldsymbol{\eta}\right\|_{L^2(\Omega)}^2 + 2\left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2. \tag{4.159}$$

For any $\boldsymbol{v}^h \in V_{\text{div}}^h$, one concludes by subtracting (4.154) from (4.153) that

$$\nu\left(\nabla\boldsymbol{\phi}^h,\nabla\boldsymbol{v}^h\right)+\mu\left(\nabla\cdot\boldsymbol{\phi}^h,\nabla\cdot\boldsymbol{v}^h\right) = -\nu\left(\nabla\boldsymbol{\eta},\nabla\boldsymbol{v}^h\right)-\mu\left(\nabla\cdot\boldsymbol{\eta},\nabla\cdot\boldsymbol{v}^h\right)+\left(\nabla\cdot\boldsymbol{v}^h,p\right).$$

Choosing $\boldsymbol{v}^h = \boldsymbol{\phi}^h$, and using that $(\nabla\cdot\boldsymbol{\phi}^h, q^h) = 0$ for any $q^h \in Q^h$, the error equation becomes for any $q^h \in Q^h$

$$\nu\left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2+\mu\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 = -\nu(\nabla\boldsymbol{\eta},\nabla\boldsymbol{\phi}^h)-\mu(\nabla\cdot\boldsymbol{\eta},\nabla\cdot\boldsymbol{\phi}^h)+(\nabla\cdot\boldsymbol{\phi}^h,p-q^h).$$

Applying the Cauchy–Schwarz inequality (A.10) and Young's inequality on the right-hand side and absorbing the terms with $\boldsymbol{\phi}^h$ in the left-hand side, one gets

$$\begin{aligned}&\nu\left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2+\mu\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2\\&\leq \nu\left\|\nabla\boldsymbol{\eta}\right\|_{L^2(\Omega)}^2+\mu\left\|\nabla\cdot\boldsymbol{\eta}\right\|_{L^2(\Omega)}^2+2\left\|p-q^h\right\|_{L^2(\Omega)}\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}.\end{aligned} \tag{4.160}$$

The last term on the right-hand side can be estimated also with Young's inequality

$$2\left\|p-q^h\right\|_{L^2(\Omega)}\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \leq \mu^{-1}\left\|p-q^h\right\|_{L^2(\Omega)}^2+\mu\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2, \tag{4.161}$$

which leads to

$$\left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 \leq \left\|\nabla\boldsymbol{\eta}\right\|_{L^2(\Omega)}^2+\frac{\mu}{\nu}\left\|\nabla\cdot\boldsymbol{\eta}\right\|_{L^2(\Omega)}^2+\frac{1}{\mu\nu}\inf_{q^h\in Q^h}\left\|p-q^h\right\|_{L^2(\Omega)}^2.$$

Finally, (4.159) gives

$$\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}^2 \leq 4\left\|\nabla\boldsymbol{\eta}\right\|_{L^2(\Omega)}^2+2\frac{\mu}{\nu}\left\|\nabla\cdot\boldsymbol{\eta}\right\|_{L^2(\Omega)}^2+\frac{2}{\mu\nu}\inf_{q^h\in Q^h}\left\|p-q^h\right\|_{L^2(\Omega)}^2$$

for all $\boldsymbol{v}^h \in V_{\text{div}}^h$, which is just the statement of the theorem. ■

Remark 4.123 (On Theorem 4.122) In contrast to the proof of Theorem 4.21, the norm of the divergence of the test function was not estimated by the norm of the gradient, see (4.24). Thus, the best approximation error with respect to the divergence appears in estimate (4.158). Now, the consequences of the error bound (4.158) on the choice of μ can be studied for two different cases. These two cases are characterized by whether or not the sequence of divergence-free subspaces of the velocity space has optimal approximation property. □

Corollary 4.124 (Application to Taylor–Hood Pairs of Finite Elements) *Consider $V^h/Q^h = P_k/P_{k-1}$, $k \geq 2$, on a family of quasi-uniform triangulations and assume that for the solution of (4.153) it holds $(\boldsymbol{u}, p) \in H^{k+1}(\Omega) \times H^k(\Omega)$.*

i) If $\left\{V^h_{\text{div,div}}\right\}$ does not possess the optimal approximation property, then the a-priori estimate (4.158) has the form

$$\left\| \nabla \left(\boldsymbol{u} - \boldsymbol{u}^h\right)\right\|^2_{L^2(\Omega)} \leq \left(4 + \frac{2\mu}{\nu}\right) C^2_{V^h_{\text{div}}} h^{2k} \left\|\boldsymbol{u}\right\|^2_{H^{k+1}(\Omega)} + \frac{2C^2_{Q^h}}{\mu\nu} h^{2k} \left\|p\right\|^2_{H^k(\Omega)}. \tag{4.162}$$

ii) If $\left\{V^h_{\text{div,div}}\right\}$ has the optimal approximation property, one obtains the a-priori error estimate

$$\left\| \nabla \left(\boldsymbol{u} - \boldsymbol{u}^h\right)\right\|^2_{L^2(\Omega)} \leq \min\left\{\left(4 + \frac{2\mu}{\nu}\right) C^2_{V^h_{\text{div}}}, 4C^2_{V^h_{\text{div,div}}}\right\} h^{2k} \left\|\boldsymbol{u}\right\|^2_{H^{k+1}(\Omega)} + \frac{2C^2_{Q^h}}{\mu\nu} h^{2k} \left\|p\right\|^2_{H^k(\Omega)}. \tag{4.163}$$

The constants $C_{Q^h}, C_{V^h_{\text{div,div}}}, C_{V^h_{\text{div}}}$ are constants coming from interpolation estimates, where $C_{V^h_{\text{div,div}}}$ and $C_{V^h_{\text{div}}}$ depend either on the inverse of the discrete inf-sup constant β^h_{is} or on the inverse of local inf-sup constants, compare Remark 4.29.

Proof

i) For this case, one can only use (3.41), i.e., that

$$\left\|\nabla \cdot \boldsymbol{v}^h\right\|_{L^2(\Omega)} = \left\|\nabla \cdot (\boldsymbol{u} - \boldsymbol{v}^h)\right\|_{L^2(\Omega)} \leq \left\|\nabla(\boldsymbol{u} - \boldsymbol{v}^h)\right\|_{L^2(\Omega)} \tag{4.164}$$

holds in this setting. Then, one applies for P_2/P_1 in three dimensions (3.65), for all other pairs (3.71), and the interpolation error estimate (C.14) to prove (4.162).

ii) In this case, one gets a second estimate besides (4.164), since one can choose $\boldsymbol{v}^h \in V^h_{\text{div,div}}$ in (4.158) and apply estimate (4.157). Note that choosing a special function yields an upper bound of the infimum. Hence, the velocity error term can be also bounded, using (3.65) for P_2/P_1 in three dimensions, (3.71) for all other pairs, and the interpolation error estimate (C.14), by

$$\inf_{\boldsymbol{v}^h \in V^h_{\text{div}}} \left(4\left\|\nabla\left(\boldsymbol{u} - \boldsymbol{v}^h\right)\right\|^2_{L^2(\Omega)} + 2\frac{\mu}{\nu}\left\|\nabla \cdot \boldsymbol{v}^h\right\|^2_{L^2(\Omega)}\right) \leq 4C^2_{V^h_{\text{div,div}}} h^{2k} \left\|\boldsymbol{u}\right\|^2_{H^{k+1}(\Omega)},$$

since $\left\|\nabla \cdot \boldsymbol{v}^h\right\|^2_{L^2(\Omega)}$ vanishes. Combining both possible estimates gives (4.163).

■

Remark 4.125 (Good Choices for the Stabilization Parameters in the Case of the Taylor–Hood Finite Element) The two cases from Corollary 4.124 will be discussed now in more detail.

i) If $\left\{V^h_{\mathrm{div,div}}\right\}$ does not have an optimal approximation property, one can consider the right-hand side of (4.162) as a function depending on μ. This function has a minimum which can be determined by elementary calculus, namely by checking the necessary and a sufficient condition for a local minimum. One obtains

$$\mu_{\mathrm{opt}} \approx \frac{C_{Q^h}}{C_{V^h_{\mathrm{div}}}} \frac{\|p\|_{H^k(\Omega)}}{\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}}. \tag{4.165}$$

Hence, with respect to ν and h, the parameter choice $\mu = \mathcal{O}(1)$ is deduced. However, it should be emphasized that μ_{opt} from (4.165) may be quite large, whenever the velocity norm is small compared with the pressure norm and that this situation can happen in practice.

Inserting μ_{opt} into the error estimate (4.162) gives

$$\begin{aligned} &\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)} \\ &\le 2h^k\left(C^2_{V^h_{\mathrm{div}}}\|\boldsymbol{u}\|^2_{H^{k+1}(\Omega)} + \frac{1}{\nu}C_{V^h_{\mathrm{div}}}C_{Q^h}\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}\|p\|_{H^k(\Omega)}\right)^{1/2}. \end{aligned} \tag{4.166}$$

This estimate reveals a direct dependency of the error of the gradient of the velocity on the pressure of the form $\nu^{-1/2}\|p\|^{1/2}_{H^k(\Omega)}$, even for the best possible stabilization parameter. On the one hand, this dependency on the viscosity is weaker than for the Galerkin discretization, where the error bound depends on ν^{-1}, see (4.45), but on the other hand, the bound of the velocity error is still influenced by the inverse of the viscosity.

ii) If $\left\{V^h_{\mathrm{div,div}}\right\}$ has the optimal approximation property, the right-hand side of estimate (4.163) is not as easy to analyze. Numerical evidence shows, see Jenkins et al. (2014), that, depending on the complexity of the pressure, there may or there may be not an optimal μ, since for $\|p\|_{H^k(\Omega)} \gg \|\boldsymbol{u}\|_{H^{k+1}(\Omega)}$ one has $\mu_{\mathrm{opt}} = \infty$, which is not feasible in practice. Therefore, giving up the idea of finding the optimal μ, one wants to find a good μ, which should not be infinity. To this end, one can use the criterion that the contribution of the pressure in the error bound (4.163) equals the maximum possible contribution of the velocity $4C^2_{V^h_{\mathrm{div,div}}} h^{2k}\|\boldsymbol{u}\|^2_{H^{k+1}(\Omega)}$, which is already asymptotically optimal. This criterion leads to

$$\mu_{\mathrm{good}} \approx \frac{1}{2\nu}\left(\frac{C_{Q^h}}{C_{V^h_{\mathrm{div,div}}}} \frac{\|p\|_{H^k(\Omega)}}{\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}}\right)^2. \tag{4.167}$$

Numerical studies in Jenkins et al. (2014) show that this value gives in fact good results. It is interesting that only in the second case μ_{good} is depending on ν, which could be observed in the numerical studies of Jenkins et al. (2014) as well. Inserting μ_{good} in (4.163) gives the error estimate

$$\left\| \nabla \left(\boldsymbol{u} - \boldsymbol{u}^h \right) \right\|_{L^2(\Omega)} \leq \sqrt{8} C_{V^h_{\text{div,div}}} h^k \left\| \boldsymbol{u} \right\|_{H^{k+1}(\Omega)}, \tag{4.168}$$

which does not directly depend on ν and p. But of course $\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}$ might still depend on ν. If $\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}$ does not depend on ν, (4.168) shows that the grad-div stabilization is able to deliver optimal uniform approximations, if there exists a sequence of subspaces of optimally converging divergence-free finite element functions.

Both proposed stabilization parameters (4.165) and (4.167) do not depend on the mesh width h. □

Corollary 4.126 (Application to the MINI Element) *Let $V^h/Q^h = P_1^{\text{bubble}}/P_1$ on a family of quasi-uniform meshes and assume that for the solution of* (4.153) *it holds* $(\boldsymbol{u}, p) \in H^2(\Omega) \times H^2(\Omega)$.

i) If $\left\{V^h_{\text{div,div}}\right\}$ does not possess optimal approximation property, then the a-priori estimate (4.158) *has the form*

$$\left\| \nabla \left(\boldsymbol{u} - \boldsymbol{u}^h \right) \right\|^2_{L^2(\Omega)} \leq \left(4 + \frac{2\mu}{\nu} \right) C^2_{V^h_{\text{div}}} h^2 \left\| \boldsymbol{u} \right\|^2_{H^2(\Omega)} + \frac{2C^2_{Q^h}}{\mu\nu} h^4 \left\| p \right\|^2_{H^2(\Omega)}. \tag{4.169}$$

ii) If $\left\{V^h_{\text{div,div}}\right\}$ has optimal approximation property, one obtains the a-priori error estimate

$$\begin{aligned} \left\| \nabla \left(\boldsymbol{u} - \boldsymbol{u}^h \right) \right\|^2_{L^2(\Omega)} \leq {} & \min \left\{ \left(4 + \frac{2\mu}{\nu} \right) C^2_{V^h_{\text{div}}}, 4C^2_{V^h_{\text{div,div}}} \right\} h^2 \left\| \boldsymbol{u} \right\|^2_{H^2(\Omega)} \\ & + \frac{2C^2_{Q^h}}{\mu\nu} h^4 \left\| p \right\|^2_{H^2(\Omega)}. \end{aligned} \tag{4.170}$$

The constants $C_{Q^h}, C_{V^h_{\text{div,div}}}, C_{V^h_{\text{div}}}$ are constants arising from interpolation estimates, where $C_{V^h_{\text{div,div}}}$ and $C_{V^h_{\text{div}}}$ depend on the inverse of β^h_{is}.

Proof The proof is performed analogously as the proof of Corollary 4.124. Note that for the MINI element the best approximation properties are already obtained for $(\boldsymbol{u}, p) \in H^2(\Omega) \times H^2(\Omega)$ and that there is no improvement if a higher regularity of the solution is assumed. In addition, one has to observe that (3.65) has to be applied for estimating the best approximation error in V^h_{div} since the MINI element does not satisfy the assumptions of estimate (3.71). ■

Remark 4.127 (Good Choices for the Stabilization Parameters in the Case of the MINI Element) The two cases from Corollary 4.126 can be discussed in the same way as for the Taylor–Hood pair of spaces.

i) If $\{V^h_{\mathrm{div,div}}\}$ does not have an optimal approximation property, one can compute an optimal value for the stabilization parameter from the right-hand side of (4.169)

$$\mu_{\mathrm{opt}} \approx h \frac{C_{Q^h}}{C_{V^h_{\mathrm{div}}}} \frac{\|p\|_{H^2(\Omega)}}{\|\boldsymbol{u}\|_{H^2(\Omega)}}.$$

ii) If $\{V^h_{\mathrm{div,div}}\}$ possesses the optimal approximation property, the second velocity term and the pressure term on the right-hand side of (4.170) can be equilibrated, leading to

$$\mu_{\mathrm{good}} \approx \frac{h^2}{2\nu} \left(\frac{C_{Q^h}}{C_{V^h_{\mathrm{div,div}}}} \frac{\|p\|_{H^2(\Omega)}}{\|\boldsymbol{u}\|_{H^2(\Omega)}} \right)^2 .$$

In contrast to the Taylor–Hood pair of spaces, good stabilization parameters for the MINI element depend on the mesh width h. □

Remark 4.128 (On the Grad-Div Stabilization)

- Since $V^h_{\mathrm{div,div}} \subset V^h_{\mathrm{div}}$, it can be expected that $C_{V^h_{\mathrm{div,div}}}$ is larger than $C_{V^h_{\mathrm{div}}}$.
- The dependency of the stabilization parameter on higher order norms of the solution was derived in Olshanskii et al. (2009), Heister and Rapin (2013), Jenkins et al. (2014).
- Optimal stabilization parameters on the basis of the $L^2(\Omega)$ error estimate of the pressure were also derived in Jenkins et al. (2014). These parameters differ from the parameters obtained from the estimate for the $L^2(\Omega)$ error of the velocity gradient.
- In Jenkins et al. (2014), comprehensive numerical studies were performed to support the analytical results with respect to the dependency of a good stabilization parameter on ν, h, and the ratio of the norm of the pressure and the velocity. It was shown that the conclusions from the analysis are well reflected in the numerical results.
- The grad-div stabilization will be studied also in the context of the Oseen equations, see Sect. 5.3.2. It turns out that the asymptotic optimal parameter choice differs for inf-sup stable pairs of finite element spaces and for equal order pairs where the PSPG method is used for stabilizing the violation of the discrete inf-sup condition, compare Remark 5.42.
- A grad-div stabilization term arises also in modeling turbulent flows with some variational multiscale (VMS) methods, see Remarks 8.226 and 8.244.
- It is reported in the literature that introducing the grad-div stabilization might improve the performance of solvers for the arising discrete system, for details

see Glowinski and Le Tallec (1989), Vassilevski and Lazarov (1996), Bychenkov and Chizonkov (1999), Heister and Rapin (2013).

□

Remark 4.129 (Implementation) The grad-div term affects only the velocity-velocity coupling. The matrix entries have the form

$$\left(\tilde{A}_{kl}\right)_{ij} = \left(\partial_l \phi_j^h, \partial_k \phi_i^h\right), \quad k, l = 1, \ldots, d, \; i, j = 1, \ldots, N_v,$$

such that the matrix A is of the following structure

$$\tilde{A} = \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} & \tilde{A}_{13} \\ \tilde{A}_{12}^T & \tilde{A}_{22} & \tilde{A}_{23} \\ \tilde{A}_{13}^T & \tilde{A}_{23}^T & \tilde{A}_{33} \end{pmatrix}.$$

The symmetry of the off-diagonal blocks of $\tilde{A}$ follows directly from the symmetry of the grad-div term.

The matrix $\tilde{A}$ has to be added to the matrix A which represents the discretization of the viscous term. The sum of both matrices, again denoted by A, forms the velocity-velocity coupling in the linear saddle point problem. Thus, the use of the grad-div stabilization prevents the velocity-velocity matrix A from being a block-diagonal matrix of the form (4.93). In particular, if the gradient form of the viscous term is used, then the grad-div stabilization introduces additional matrix blocks. A grad-div stabilization with reduced memory requirements, a so-called sparse grad-div stabilization, was proposed in Bowers et al. (2014). □

Example 4.130 (Analytical Solution, Continuation of Example 4.31) Consider the Stokes equations (4.44) with $\nu = 1$, the solution given by Example D.3, the Taylor–Hood pair of finite element spaces P_2/P_1, and the irregular grid depicted in Fig. 4.2. Thus, there is the same setup as in Example 4.31. Here, only the effect of the grad-div stabilization on the violation of the incompressibility constraint will be illustrated.

The application of the grad-div stabilization requires the choice of the stabilization parameters $\{\mu_K\}$. For the irregular grid from Fig. 4.2 one would not expect that there exists a sequence of divergence-free subspaces with optimal convergence property. Thus, with respect to the bound of the $L^2(\Omega)$ error of the velocity gradient, (4.165) for $k = 2$ would be the optimal parameter. However, in practice the solution is unknown and consequently, norms of the solution are not available. For inf-sup stable pairs of finite element spaces, one usually uses constant stabilization parameters of the order of unity, e.g., $\mu \in [0.1, 10]$. In Olshanskii (2002), good results were obtained with $\mu = 0.2$.

Figure 4.18 shows the effect of applying the grad-div stabilization with different parameters on $\left\|\nabla \cdot \boldsymbol{u}^h\right\|_{L^2(\Omega)}$. It can be seen that there is some improvement concerning the violation of the incompressibility constraint on coarse grids. But on finer grids, this effect becomes negligible. □

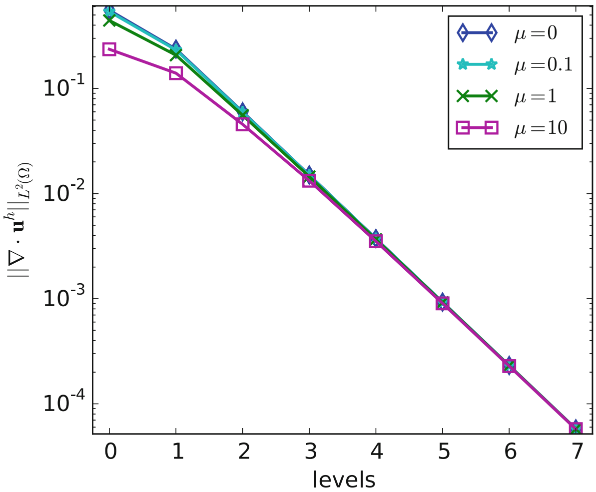

Fig. 4.18 Example 4.130. Convergence of the errors $\|\nabla \cdot \boldsymbol{u}^h\|_{L^2(\Omega)}$ for P_2/P_1 with grad-div stabilization and different values of the stabilization parameter

Remark 4.131 (Summary) The grad-div stabilization is a simple and popular stabilization for finite element discretizations of equations modeling incompressible flow problems. Compared with the Galerkin discretization, the impact of the inverse of the viscosity and of the pressure on error bounds for the velocity becomes weaker. However, the grad-div stabilization generally neither removes this impact nor leads to (weakly) divergence-free discrete solutions. □

4.6.2 Choosing Appropriate Test Functions

Remark 4.132 (Goal and Idea) The principle idea of this approach can be explained easiest for the Crouzeix–Raviart finite element P_1^{nc}/P_0. In fact, its first presentation in the literature in Linke (2014) was for this pair of finite element spaces.

A finite element error estimate for the scaled Stokes equations (4.44) and the velocity error $\|\boldsymbol{u} - \boldsymbol{u}^h\|_{V^h}$ is given in (4.86). A large norm of the pressure gradient or a small viscosity will have a large impact on the error bound. This impact can be seen in actual simulations, see Example 4.65. The goal of the approach proposed in Linke (2014) consists in deriving a method such that the pressure term in (4.86) vanishes and with that also the dependency of the error bound on the viscosity. Methods with this property are called pressure-robust.

A detailed inspection of the finite element error analysis for the Crouzeix–Raviart finite element shows that the pressure term appears in the estimate only because the term (4.70) does not vanish for functions $\boldsymbol{v} \in V_{\mathrm{div}}^h$. The term (4.70) appears on the right-hand side of (4.65) coming from the term $(\boldsymbol{f}, \boldsymbol{v})_K$. The principal idea in

Linke (2014) consists in using on the right-hand side of the Stokes equations a test function such that (4.70) vanishes. To this end, the function $\boldsymbol{v}$ has to be chosen such that $\boldsymbol{v} \in V_{\text{div}} \oplus V^h_{\text{Hdiv}}$, where V^h_{Hdiv} is an appropriate space. It has to be a subspace of $H_{\text{div}}(\Omega)$, see (3.39) and in particular a subspace of $H(\text{div}, \Omega)$. The conditions for the last property are given in Lemma 3.66. The perhaps simplest space which satisfies these conditions is the Raviart–Thomas space of lowest order RT_0, see Example B.45. In Linke (2014) it was proposed to use on the right-hand side of the Stokes equations a test function from RT_0, which is an appropriate projection of the Crouzeix–Raviart test function. □

Remark 4.133 (Interpolation Operators) The definition and the analysis of the method requires the use of two interpolation operators.

First, the interpolation operator $P^h_E : V \to V^h = P_1^{\text{nc}}$ defined in (3.143) is used.

The second interpolation operator is given by $P^h_{E,\text{RT}_0} : V \cup V^h \to \text{RT}_0$ with

$$\left(P^h_{E,\text{RT}_0}\boldsymbol{v} \cdot \boldsymbol{n}_E\right)(\boldsymbol{m}_E) = \begin{cases} \dfrac{1}{|E|}\displaystyle\int_E \boldsymbol{v} \cdot \boldsymbol{n}_E \, ds & \text{if } E \in \mathcal{E}^h, \\ 0 & \text{if } E \in \overline{\mathcal{E}}^h \setminus \mathcal{E}^h. \end{cases} \tag{4.171}$$

Note that (4.171) defines constant values for the normal component on each face E. With that, the degrees of freedom for a RT_0 function are determined.

From Acosta and Durán (1999), the following interpolation estimate is known

$$\left\|\boldsymbol{v} - P^h_{E,\text{RT}_0}\boldsymbol{v}\right\|_{L^2(\Omega)} \le Ch \left\|\boldsymbol{v}\right\|_{V^h} \quad \forall\ \boldsymbol{v} \in V \cup V^h, \tag{4.172}$$

where the constant depends on the maximal angle of the mesh cells and $\|\cdot\|_{V^h}$ is defined in (3.54).

The divergence of a function from RT_0 is a piecewise constant function. Equipping RT_0 with homogeneous normal components on the boundary of Ω, integration by parts gives

$$\int_\Omega \nabla \cdot \boldsymbol{v} \, d\boldsymbol{x} = 0, \quad \forall\ \boldsymbol{v} \in \text{RT}_0,$$

such that the divergence of a function from RT_0 belongs to Q^h. □

Lemma 4.134 (The Divergence of the Projections) *Denote by $P^h_{L^2}$ the $L^2(\Omega)$ projection onto Q^h. Then it is*

$$\text{div}^h\left(P^h_E\boldsymbol{v}\right) = P^h_{L^2}(\nabla \cdot \boldsymbol{v}) \quad \forall\ \boldsymbol{v} \in V \tag{4.173}$$

and

$$\nabla \cdot \left(P^h_{E,\text{RT}_0}\boldsymbol{v}\right) = P^h_{L^2}(\nabla \cdot \boldsymbol{v}) \quad \forall\ \boldsymbol{v} \in V \cup V^h. \tag{4.174}$$

In addition, it is

$$P^h_{E,\mathrm{RT}_0}\boldsymbol{v} \in H_{\mathrm{div}}(\Omega) \quad \forall\ \boldsymbol{v} \in V_{\mathrm{div}} \cup V^h_{\mathrm{div}}, \tag{4.175}$$

i.e., $P^h_{E,\mathrm{RT}_0}\boldsymbol{v}$ *weakly divergence-free.*

Proof Relation (3.144) can be written in the form

$$-\sum_{K\in\mathcal{T}^h}\left(\nabla\cdot P^h_E\boldsymbol{v}, q^h\right)_K = -\sum_{K\in\mathcal{T}^h}\left(\nabla\cdot\boldsymbol{v}, q^h\right)_K \quad \forall\ \boldsymbol{v}\in V,\ \forall\ q^h\in Q^h,$$

which is exactly the statement (4.173).

The relation (4.174) is derived exactly in the same form as (3.144).

Using integration by parts, replacing the integrals on the faces exactly by the mid point rule, which is possible because the functions are constant on the faces, using definition (4.171), and applying again integration by parts gives

$$\begin{aligned}\int_K \nabla\cdot\left(P^h_{E,\mathrm{RT}_0}\boldsymbol{v}\right)\,d\boldsymbol{x} &= \sum_{E\subset\partial K}\int_E \left(P^h_{E,\mathrm{RT}_0}\boldsymbol{v}\right)\cdot\boldsymbol{n}_E\,ds = \sum_{E\subset\partial K}|E|\left(P^h_{E,\mathrm{RT}_0}\boldsymbol{v}\cdot\boldsymbol{n}_E\right)(\boldsymbol{m}_E)\\ &= \sum_{E\subset\partial K}\int_E \boldsymbol{v}\cdot\boldsymbol{n}_E\,ds = \int_K\nabla\cdot\boldsymbol{v}\,d\boldsymbol{x}\end{aligned} \tag{4.176}$$

for an arbitrary mesh cell $K\in\mathcal{T}^h$. If

$$\int_K \nabla\cdot\boldsymbol{v}\,d\boldsymbol{x} = 0, \tag{4.177}$$

one gets, using that the divergence of a function from $\mathrm{RT}_0(K)$ is constant,

$$0 = \int_K\nabla\cdot\boldsymbol{v}\,d\boldsymbol{x} = |K|\ \nabla\cdot\left(P^h_{E,\mathrm{RT}_0}\boldsymbol{v}\right)\Big|_K.$$

Since $|K| > 0$, it follows that $\nabla\cdot\left(P^h_{E,\mathrm{RT}_0}\boldsymbol{v}\right)\big|_K = 0$. Condition (4.177) is clearly satisfied for functions $\boldsymbol{v}\in V_{\mathrm{div}}$. From Lemma 4.62, it follows that also functions $\boldsymbol{v}\in V^h_{\mathrm{div}}$ fulfill (4.177). Since $P^h_{E,\mathrm{RT}_0}\boldsymbol{v}\in H(\mathrm{div},\Omega)$, it follows that $P^h_{E,\mathrm{RT}_0}\boldsymbol{v}\in H_{\mathrm{div}}(\Omega)$. ∎

Remark 4.135 (The Modified Problem) The modified problem reads as follows: Find $\left(\boldsymbol{u}^h, q^h\right)\in V^h\times Q^h = P_1^{\mathrm{nc}}\times P_0$ such that

$$\begin{aligned}\nu a^h\left(\boldsymbol{u}^h,\boldsymbol{v}^h\right) + b^h\left(\boldsymbol{v}^h,p^h\right) &= \left(\boldsymbol{f}, P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h\right) && \forall\ \boldsymbol{v}^h\in V^h,\\ b^h\left(\boldsymbol{u}^h,q^h\right) &= 0 && \forall\ q^h\in Q^h,\end{aligned} \tag{4.178}$$

with the bilinear forms given in (4.18). In comparison with the original problem (4.59), apart of the viscosity, only the projection of the test function on the right-hand side is introduced. □

Lemma 4.136 (Consistency Error Estimate) *Let* $(\boldsymbol{u}, p)$ *be a sufficiently smooth solution of the Stokes equations* (4.2) *with* $\boldsymbol{u} \in C^1\left(\overline{\Omega}\right) \cap V$, $p \in C\left(\overline{\Omega}\right) \cap Q$. *Consider a family of quasi-uniform triangulations, then it holds for all* $\boldsymbol{v} \in V_{\mathrm{div}} \oplus V_{\mathrm{div}}^h$

$$\left|\nu a^h(\boldsymbol{u}, \boldsymbol{v}) - \left(\boldsymbol{f}, P_{E,\mathrm{RT}_0}^h \boldsymbol{v}\right)\right| \le C\nu h\, |\boldsymbol{u}|_{H^2(\Omega)}\, \|\boldsymbol{v}\|_{V^h}\,. \tag{4.179}$$

Proof The proof starts exactly like the proof of Lemma 4.55.

Let $\boldsymbol{v} \in V_{\mathrm{div}} \oplus V_{\mathrm{div}}^h$ be arbitrary. Using the momentum equation of the Stokes problem (4.1) yields

$$\begin{aligned}
&\nu a^h(\boldsymbol{u}, \boldsymbol{v}) - \left(\boldsymbol{f}, P_{E,\mathrm{RT}_0}^h \boldsymbol{v}\right) \\
&= \sum_{K \in \mathcal{T}^h} \nu\, (\nabla \boldsymbol{u}, \nabla \boldsymbol{v})_K - \left(\boldsymbol{f}, P_{E,\mathrm{RT}_0}^h \boldsymbol{v}\right)_K \\
&= \sum_{K \in \mathcal{T}^h} \nu\, (\nabla \boldsymbol{u}, \nabla \boldsymbol{v})_K - (-\nu \Delta \boldsymbol{u}, \boldsymbol{v})_K + \sum_{K \in \mathcal{T}^h} \left(-\nu \Delta \boldsymbol{u}, \boldsymbol{v} - P_{E,\mathrm{RT}_0}^h \boldsymbol{v}\right)_K \\
&\quad - \sum_{K \in \mathcal{T}^h} \left(\nabla p, P_{E,\mathrm{RT}_0}^h \boldsymbol{v}\right)_K .
\end{aligned} \tag{4.180}$$

The first sum in (4.180) was already estimated in (4.69) for $\nu = 1$. Applying the same techniques gives

$$\sum_{K \in \mathcal{T}^h} \nu\, (\nabla \boldsymbol{u}, \nabla \boldsymbol{v})_K + (\nu \Delta \boldsymbol{u}, \boldsymbol{v})_K \le C\nu h\, |\boldsymbol{u}|_{H^2(\Omega)}\, \|\boldsymbol{v}\|_{V^h}\,. \tag{4.181}$$

The second sum is estimated with the Cauchy–Schwarz inequality (A.10), the Cauchy–Schwarz inequality for sums (A.2), and the interpolation estimate (4.172)

$$\begin{aligned}
\sum_{K \in \mathcal{T}^h} \left(-\nu \Delta \boldsymbol{u}, \boldsymbol{v} - P_{E,\mathrm{RT}_0}^h \boldsymbol{v}\right)_K &\le \sum_{K \in \mathcal{T}^h} \nu\, \|\Delta \boldsymbol{u}\|_{L^2(K)} \left\|\boldsymbol{v} - P_{E,\mathrm{RT}_0}^h \boldsymbol{v}\right\|_{L^2(K)} \\
&\le \nu\, \|\Delta \boldsymbol{u}\|_{L^2(\Omega)} \left\|\boldsymbol{v} - P_{E,\mathrm{RT}_0}^h \boldsymbol{v}\right\|_{L^2(\Omega)} \\
&\le C\nu h\, |\boldsymbol{u}|_{H^2(\Omega)}\, \|\boldsymbol{v}\|_{V^h}\,.
\end{aligned} \tag{4.182}$$

Note that for estimating the first and second sum on the right-hand side of (4.180) the property of $\boldsymbol{v}$ being (discretely) divergence-free is not needed, compare also Remark 4.57.

Finally, integration by parts is applied for the last sum of (4.180)

$$-\sum_{K\in\mathcal{T}^h}\left(\nabla p, P^h_{E,\mathrm{RT}_0}\boldsymbol{v}\right)_K = \sum_{K\in\mathcal{T}^h}\left(\nabla\cdot\left(P^h_{E,\mathrm{RT}_0}\boldsymbol{v}\right), p\right)_K - \int_{\partial K} pP^h_{E,\mathrm{RT}_0}\boldsymbol{v}\cdot\boldsymbol{n}_{\partial K}\,ds.$$

The first term on the right-hand side vanishes since $P^h_{E,\mathrm{RT}_0}\boldsymbol{v}$ is divergence-free, see (4.175). The integrals on the interior faces vanish because p and $P^h_{E,\mathrm{RT}_0}\boldsymbol{v}\cdot\boldsymbol{n}_{\partial K}$ are both continuous. On the boundary faces, $P^h_{E,\mathrm{RT}_0}\boldsymbol{v}\cdot\boldsymbol{n}_{\partial K} = 0$, see (4.171). Altogether, the last sum of (4.180) is zero.

Collecting all estimates proves the statement of the lemma. ■

Theorem 4.137 (Error Estimate for the V^h Norm of the Velocity and the $L^2(\Omega)$ Norm of the Pressure) *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, be a bounded domain with polyhedral and Lipschitz continuous boundary and let $(\boldsymbol{u},p)$, the unique solution of the scaled Stokes problem (4.44), be sufficiently smooth in the sense of Remark 4.56. Consider a quasi-uniform family of triangulations and let $\left(\boldsymbol{u}^h,p^h\right)\in P_1^{\mathrm{nc}}\times P_0$ be the unique solution of the finite element problem (4.178) with modified test function on the right-hand side. Then the following error estimates are valid*

$$\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{V^h} \le Ch\,|\boldsymbol{u}|_{H^2(\Omega)}\,, \tag{4.183}$$

$$\left\|p-p^h\right\|_{L^2(\Omega)} \le Ch\left(1+\frac{1}{\beta^h_{\mathrm{is}}}\right)\left(\nu\,|\boldsymbol{u}|_{H^2(\Omega)}+\|\nabla p\|_{L^2(\Omega)}\right). \tag{4.184}$$

Proof Exactly as in the proof of Lemma 4.51 (second lemma of Strang), one derives for the scaled Stokes problem (4.44) and the discretization (4.178) the abstract error estimate

$$\begin{aligned}\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{V^h} \le{}& 2\inf_{\boldsymbol{v}^h\in V^h_{\mathrm{div}}}\left\|\boldsymbol{u}-\boldsymbol{v}^h\right\|_{V^h}\\ &+\frac{1}{\nu}\inf_{\boldsymbol{v}^h\in V^h_{\mathrm{div}},\|\boldsymbol{v}^h\|_{V^h}=1}\left|\nu a^h\left(\boldsymbol{u},\boldsymbol{v}^h\right)-\left(\boldsymbol{f},P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h\right)\right|.\end{aligned}$$

Inserting the best approximation error estimate (4.63) and the consistency error estimate (4.179) in this abstract estimate proves the velocity estimate (4.183).

The estimate of the pressure error proceeds along the way of the proof of Theorem 4.60. The first steps, until estimate (4.79) are exactly the same. The estimate of the next term starts as follows

$$\begin{aligned} b^h\left(\boldsymbol{v}^h,p^h-p\right) &= -\nu a^h\left(\boldsymbol{u}^h,\boldsymbol{v}^h\right)+\left(\boldsymbol{f},P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h\right)-b^h\left(\boldsymbol{v}^h,p\right) \\ &= \nu a^h\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{v}^h\right)-\nu a^h\left(\boldsymbol{u},\boldsymbol{v}^h\right)+\left(\boldsymbol{f},P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h\right)-b^h\left(\boldsymbol{v}^h,p\right). \end{aligned} \tag{4.185}$$

Using the Cauchy–Schwarz inequality (A.10) and the error bound (4.183) for the velocity gives

$$\nu a^h\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{v}^h\right) \le \nu \left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{V^h} \left\|\boldsymbol{v}^h\right\|_{V^h} \le C\nu h \left|\boldsymbol{u}\right|_{H^2(\Omega)} \left\|\boldsymbol{v}^h\right\|_{V^h}.$$

The other terms on the right-hand side of (4.185) are written in the form

$$\begin{aligned}
&-\nu a^h\left(\boldsymbol{u},\boldsymbol{v}^h\right)+\left(\boldsymbol{f},P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h\right)-b^h\left(\boldsymbol{v}^h,p\right)\\
&=\sum_{K\in\mathcal{T}^h}\left[-\nu\left(\nabla\boldsymbol{u},\nabla\boldsymbol{v}^h\right)_K+\left(-\nu\Delta\boldsymbol{u},P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h\right)_K+\left(\nabla p,P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h\right)_K+\left(\nabla\cdot\boldsymbol{v}^h,p\right)_K\right].
\end{aligned}$$

The first two terms were estimated in (4.181), (4.182), noting that for these estimates the discrete divergence-free condition is not necessary, leading to

$$\sum_{K\in\mathcal{T}^h}\left[-\nu\left(\nabla\boldsymbol{u},\nabla\boldsymbol{v}^h\right)_K+\left(-\nu\Delta\boldsymbol{u},P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h\right)_K\right]\le C\nu h\left|\boldsymbol{u}\right|_{H^2(\Omega)}\left\|\boldsymbol{v}^h\right\|_{V^h}.$$

For the terms with the pressure, integration by parts is applied, the Cauchy–Schwarz inequality is used, estimate (4.172) is utilized, and the integrals on the faces are estimate with (4.71)

$$\begin{aligned}
&\sum_{K\in\mathcal{T}^h}\left[\left(\nabla p,P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h\right)_K+\left(\nabla\cdot\boldsymbol{v}^h,p\right)_K\right]\\
&=\sum_{K\in\mathcal{T}^h}\left[\left(\nabla p,P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h-\boldsymbol{v}^h\right)_K+\sum_{E\in\partial K}\int_E p\boldsymbol{v}^h\cdot\boldsymbol{n}_E\,d\boldsymbol{s}\right]\\
&\le\left\|\nabla p\right\|_{L^2(\Omega)}\left\|P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h-\boldsymbol{v}^h\right\|_{L^2(\Omega)}+Ch\left\|\nabla p\right\|_{L^2(\Omega)}\left\|\boldsymbol{v}^h\right\|_{V^h}\\
&\le Ch\left\|\nabla p\right\|_{L^2(\Omega)}\left\|\boldsymbol{v}^h\right\|_{V^h}.
\end{aligned}$$

Collecting all estimates leads in the same way as in the proof of Theorem 4.60 to estimate (4.184). ■

Remark 4.138 (On Theorem 4.137)

- The velocity solution $\boldsymbol{u}^h$ of (4.178) is not divergence-free in the sense of $H_{\mathrm{div}}(\Omega)$ since it is a function from V^h_{div}. Applying as post-processing the operator P^h_{E,RT_0} to $\boldsymbol{u}^h$ gives a divergence-free velocity field $P^h_{E,\mathrm{RT}_0}\boldsymbol{u}^h$, see (4.175).
- An optimal estimate for $\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{L^2(\Omega)}$ for the proposed method was derived in Linke et al. (2016a). The proof of this estimate requires a somewhat higher regularity of the velocity solution $\boldsymbol{u}$ than assumed in Theorem 4.137 and Remark 4.56.

□

Example 4.139 (Scaled Stokes Equations with Analytic Solution for the Modified Discretization (4.178)*)* This example continues Example 4.65, where the results for the Galerkin discretization with the Crouzeix–Raviart pair of finite element spaces are presented.

To implement the modified method (4.178), the projection $P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h$ has to be computed for each basis function $\boldsymbol{v}^h \in V^h$. Consider an arbitrary mesh cell $K \in \mathcal{T}^h$, then a function from $\mathrm{RT}_0(K)$ can be written in the form

$$\boldsymbol{v}^h_{\mathrm{RT}_0}(\boldsymbol{x}) = \boldsymbol{a} + \frac{b}{d}\left(\boldsymbol{x} - \boldsymbol{m}_K\right), \quad \boldsymbol{a} \in \mathbb{R}^d, b \in \mathbb{R}, \boldsymbol{x} \in K, \tag{4.186}$$

where $\boldsymbol{m}_K$ is the barycenter of K. To compute $P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h$, one has to determine $\boldsymbol{a}$ and b. With (4.176) and (4.186), it follows that

$$\nabla \cdot \boldsymbol{v}^h\big|_K = \nabla \cdot \left(P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h\right)\big|_K = \frac{b}{d}d = b.$$

For deriving a formula for $\boldsymbol{a}$, it is used at several places that integrals of linear functions on K and $E \subset \partial K$ can be computed exactly with the mid point rule. This argument gives

$$\int_K P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h \, d\boldsymbol{x} = |K| \left(P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h\right)(\boldsymbol{m}_K) = |K|\,\boldsymbol{a}. \tag{4.187}$$

In addition, one obtains with the product rule

$$\begin{aligned}\int_K \nabla \cdot \left(x_i P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h\right) \, d\boldsymbol{x} &= \int_K x_i \nabla \cdot P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h \, d\boldsymbol{x} + \int_K \nabla x_i \cdot P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h \, d\boldsymbol{x} \\ &= \int_K x_i \nabla \cdot P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h \, d\boldsymbol{x} + \int_K \left(P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h\right)_i \, d\boldsymbol{x},\end{aligned}$$

$i = 1, \ldots, d$. Inserting (4.187), applying integration by parts, using that the normal components and the divergence of functions from $\mathrm{RT}_0(K)$ are constant, and applying (4.176) gives

$$\begin{aligned} a_i &= \frac{1}{|K|}\left(\sum_{E\subset\partial K}\int_E x_i P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h \cdot \boldsymbol{n}_E \, d\boldsymbol{s} - \int_K x_i \nabla \cdot P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h \, d\boldsymbol{x}\right) \\ &= \frac{1}{|K|}\left(\sum_{E\subset\partial K}\left(P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h\right) \cdot \boldsymbol{n}_E \int_E x_i \, d\boldsymbol{s} - \left(\nabla \cdot P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h\right)\int_K x_i \, d\boldsymbol{x}\right) \\ &= \frac{1}{|K|}\left(\sum_{E\subset\partial K}\left(P^h_{E,\mathrm{RT}_0}\boldsymbol{v}^h\right) \cdot \boldsymbol{n}_E \, |E| \, (\boldsymbol{m}_E)_i - \nabla \cdot \boldsymbol{v}^h \, |K| \, (\boldsymbol{m}_K)_i\right). \end{aligned} \tag{4.188}$$

Using integration by parts yields

$$\nabla \cdot \boldsymbol{v}^h = \frac{1}{|K|} \int_K \nabla \cdot \boldsymbol{v}^h \, d\boldsymbol{x} = \frac{1}{|K|} \sum_{E \subset \partial K} \int_E \boldsymbol{v}^h \cdot \boldsymbol{n}_E \, ds$$
$$= \frac{1}{|K|} \sum_{E \subset \partial K} |E| \, \boldsymbol{v}^h(\boldsymbol{m}_E) \cdot \boldsymbol{n}_E.$$

Inserting this relation in (4.188) and using the definition (4.171) of the projection gives

$$\boldsymbol{a} = \frac{1}{|K|} \left(\sum_{E \subset \partial K} \left(\int_E \boldsymbol{v}^h \cdot \boldsymbol{n}_E \, ds \right) \boldsymbol{m}_E - |E| \left(\boldsymbol{v}^h(\boldsymbol{m}_E) \cdot \boldsymbol{n}_E \right) \boldsymbol{m}_K \right)$$
$$= \frac{1}{|K|} \sum_{E \subset \partial K} |E| \left(\boldsymbol{v}^h(\boldsymbol{m}_E) \cdot \boldsymbol{n}_E \right) (\boldsymbol{m}_E - \boldsymbol{m}_K) \, .$$

Results obtained with the modified discretization (4.178) are presented in Fig. 4.19. They should be compared with the results in the left column of Fig. 4.10. It can be seen that for the interesting case of small values of the viscosity ν, the velocity errors are much smaller for the modified discretization. On sufficiently fine grids, these errors become independent of ν. The lines depicting the asymptotics are the same in the pictures of Fig. 4.19 as in the respective pictures of Fig. 4.10. Thus, for large values of ν, the impact of the consistency error committed by the modified method can be observed since the curves of the errors have a larger distance to the lines of the asymptotics. With respect to the pressure, the results of the standard and the modified discretization are similar. □

Remark 4.140 (Extensions of this Approach)

- The idea presented in this section was extended to higher order inf-sup stable pairs of finite element spaces with discontinuous pressure in Linke et al. (2016b). On simplicial meshes, projections to Raviart–Thomas spaces were used and on rectangular and brick-type meshes projections to Brezzi–Douglas–Marini spaces. Optimal error estimates for the $L^2(\Omega)$ norms of the (discrete) gradient of the velocity, the velocity, and the pressure were derived. The error bounds for the velocity do not depend on the pressure and on the viscosity.
- For the Taylor–Hood pair of spaces P_2/P_1, a reconstruction operator is proposed in Lederer (2016). This operator projects to a Brezzi–Douglas–Marini space and it is defined on a patch of mesh cells.
- An extension of this approach to the Navier–Stokes equations is also possible, see Brennecke et al. (2015). For the stationary Navier–Stokes equations, also a projection of the test function in the convective term has to be applied, which leads to a modification of the stiffness matrix (in contrast to the Stokes

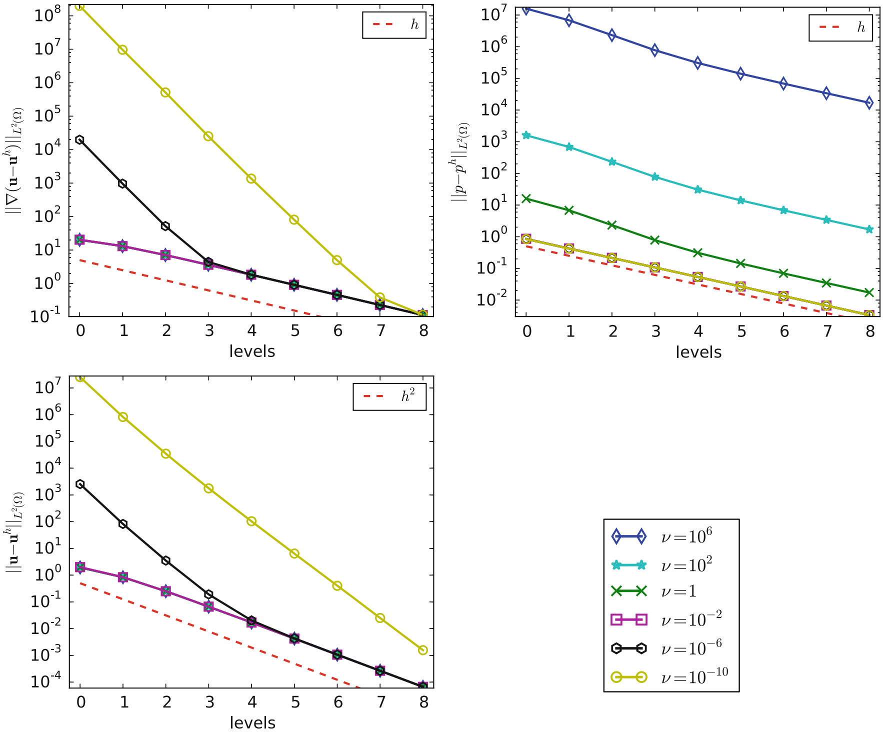

Fig. 4.19 Example 4.139. Convergence of the errors for the scaled Stokes problem and the pair of finite element spaces P_1^{nc}/P_0 with the modified discretization (4.178)

equations). In Brennecke et al. (2015), some numerical examples are presented showing the potential of this approach.

□

4.6.3 *Constructing Divergence-Free and Inf-Sup Stable Pairs of Finite Element Spaces*

Remark 4.141 (Inf-Sup Stable Pairs of Finite Element Spaces with Weakly Divergence-free Velocity Solutions) Mass conservation is ensured if the finite element velocity is divergence-free in the sense of Definition 3.32, i.e., if $\left\|\nabla \cdot \boldsymbol{u}^h\right\|_{L^2(\Omega)} = 0$. This property is surely given if $\nabla \cdot V^h \subset Q^h$. In this case, one can take the test function $q^h = \nabla \cdot \boldsymbol{u}^h$ in the discrete continuity equation which leads to

$$\left(\nabla \cdot \boldsymbol{u}^h, \nabla \cdot \boldsymbol{u}^h\right) = \left\|\nabla \cdot \boldsymbol{u}^h\right\|_{L^2(\Omega)}^2 = 0.$$

It was already discussed in Remark 3.56 that the satisfaction of the discrete inf-sup condition (3.51) and the property that the finite element velocity should be weakly divergence-free act against each other. While the former condition requires the finite element pressure space to be sufficiently small compared with the velocity space, the latter condition requires a sufficiently large space Q^h, e.g., such that $Q^h \supset \nabla \cdot V^h$.

An advantage of getting weakly divergence-free solutions was already shown in Sect. 4.2.1.2: the error bounds and the errors for the velocity do not depend on the pressure and the viscosity, see Remark 4.38 and Example 4.40.

So far only one family of inf-sup stable finite element spaces with weakly divergence-free solutions was introduced, the Scott–Vogelius family, compare Remarks 3.135 and 3.136. The most interesting pair in two dimensions, P_2/P_1^{disc}, satisfies the discrete inf-sup condition only on special grids, e.g., on barycentric-refined grids, see Fig. 3.9. The violation of the discrete inf-sup condition on general grids was shown in Example 3.73.

Since the satisfaction of the discrete inf-sup condition and the computation of weakly divergence-free finite element velocities are in some sense conflicting goals, the construction of pairs satisfying both requirements is involved. This section illustrates approaches which lead to appropriate pairs of finite element spaces. □

Remark 4.142 (The Smooth de Rham Complex or Stokes Complex in Two Dimensions) A de Rham complex is a sequence of mappings. For the Stokes problem in two dimensions, the so-called smooth de Rham complex or Stokes complex is of interest

$$\mathbb{R} \quad \to \quad H^2(\Omega) \quad \stackrel{\mathbf{curl}}{\to} \quad H^1(\Omega) \quad \stackrel{\mathrm{div}}{\to} \quad L^2(\Omega) \quad \to \quad 0, \qquad (4.189)$$

where the **curl** operator is defined in (3.153).

A de Rham complex is called to be exact if the range of each operator is the kernel of the succeeding operator. The exactness of the de Rham complex (4.189) implies:

- if $w \in H^2(\Omega)$ is curl-free, then w is constant function,
- if $\boldsymbol{v} \in H^1(\Omega)$ is divergence-free, then $\boldsymbol{v} = \mathbf{curl}\, w$ for some $w \in H^2(\Omega)$,
- the map div : $H^1(\Omega) \to L^2(\Omega)$ is surjective, since the kernel of the last operator in (4.189) is $L^2(\Omega)$.

These properties were already studied in the case of the Stokes equations. Concerning the first property, if $\mathbf{curl}\, w = 0$ in Ω, then both partial derivatives of w vanish in Ω. Hence w has to be a constant function. The second property was studied in Lemma 3.174. Note that this lemma is formulated for simply connected domains. The surjection of the divergence operator was proved in Lemma 3.43 for the setting $H_0^1(\Omega) \to L_0^2(\Omega)$. To extend this result to (4.189), one just needs to show that every constant $C \in L^2(\Omega)$ is the divergence of a function from $H^1(\Omega)$. For this purpose, one can take, e.g., $\boldsymbol{v} = (Cx, 0)^T$.

Altogether, it can be shown that the smooth de Rham complex (4.189) is exact on bounded, simply connected domains $\Omega \subset \mathbb{R}^2$ with Lipschitz boundary. □

Remark 4.143 (Finite Element Sub-complex) A finite element sub-complex of (4.189) consists of finite element spaces $W^h \subset H^2(\Omega)$, $V^h \subset H^1(\Omega)$, and $Q^h \subset L^2(\Omega)$ satisfying

$$\mathbb{R} \quad \to \quad W^h \quad \overset{\textbf{curl}}{\to} \quad V^h \quad \overset{\text{div}}{\to} \quad Q^h \quad \to \quad 0. \tag{4.190}$$

If the finite element sub-complex (4.190) is exact, then there are two important implications.

- The pair of finite element spaces V^h/Q^h satisfies the discrete inf-sup condition (3.51) provided the mapping div has a right inverse that is bounded independently of the mesh width.
- It holds div $V^h = Q^h$, i.e., using V^h/Q^h weakly divergence-free velocity fields are computed, see Remark 4.141.

□

Example 4.144 (Finite Element Sub-complex for the Scott–Vogelius Pair of Finite Element Spaces) Consider an admissible regular simplicial triangulation $\mathcal{T}^h$ of Ω consisting of mesh cells $\{K_{\text{coarse}}\}$. Each triangle is barycentric refined by connecting its vertices with its barycenter, giving a finer triangulation consisting of mesh cells $\{K_{\text{fine}}\}$.

Starting point for the derivation of the finite element sub-complex for the Scott–Vogelius pair of spaces is the Hsieh–Clough–Tocher finite element, see Clough and Tocher (1965), Ciarlet (1978, Chap. 6.1). This finite element is a composite finite element consisting of third order polynomials on each mesh cell K_{fine}, i.e., on the corresponding mesh cell K_{coarse} one has to determine three polynomials with ten degrees of freedom for each. It is required that the finite element function is continuously differentiable. Since the restriction of a cubic polynomial to an edge is a quadratic polynomial in one dimension, the satisfaction of this condition needs three degrees of freedom on each of the six edges involved in the barycentric refinement of K_{coarse}. The other degrees of freedom are the values of the function in the vertices of K_{coarse} (three degrees of freedom), the derivatives in these vertices (six degrees of freedom), and the integrals of the normal derivatives on the edges of K_{coarse} (three degrees of freedom). By the requirement on the regularity of the finite element functions, it follows that $W^h \subset H^2(\Omega)$, where W^h denotes the global Hsieh–Clough–Tocher space. Any function from W^h is uniquely determined by this requirement, the three degrees of freedom on each vertex, and the degree of freedom on each edge of the coarse mesh. Hence, one finds

$$\dim W^h = 3\#\ \text{vertices}_{\text{coarse}} + \#\ \text{edges}_{\text{coarse}}, \tag{4.191}$$

where # denotes the cardinality of a set.

Since differentiation reduces the polynomial degree by one and also the regularity by one, it follows that $V^h = \textbf{curl}\ W^h \subset H^1(\Omega)$ consists of continuous polynomials of degree two with respect to the fine mesh: $V^h = P_2$. Likewise, $Q^h = \text{div}\ V^h \subset$

$L^2(\Omega)$ consists of discontinuous linear functions with respect to the fine mesh: $Q^h = P_1^{\text{disc}}$. The pair of spaces V^h/Q^h is the Scott–Vogelius pair on the barycentric-refined mesh. It remains to show the exactness of this finite element sub-complex.

Let $\boldsymbol{v}^h \in V^h$ be weakly divergence-free. From the exactness of the smooth de Rham complex (4.189), see Remark 4.142, it can be concluded that there is a $w \in H^2(\Omega)$ with $\boldsymbol{v}^h = \textbf{curl}\ w$. By definition of V^h, one has that $\partial_x w$ and $\partial_y w$ are piecewise polynomials of degree two with respect to the fine mesh. Hence, w is a piecewise polynomial of degree three on this mesh. Since $\textbf{curl}\ w \in H^1(\Omega)$, it follows that $w \in H^2(\Omega)$. From these two properties it is concluded that $w \in W^h$.

Next, one has to show that $\text{div} : V^h \to Q^h$ is a surjection. From the considerations above, it is known that the divergence of a function from V^h is a discontinuous piecewise linear function with respect to the fine mesh, hence $\text{div}\ V^h \subset Q^h$. Since V^h and Q^h are finite-dimensional spaces, the surjection property is proved if one shows that the dimensions of $\text{div}\ V^h$ and Q^h are equal. In Q^h there are three degrees of freedom on each cell of the fine mesh, hence

$$\dim Q^h = 3\#\,\text{cells}_{\text{fine}} = 9\#\,\text{cells}_{\text{coarse}}. \tag{4.192}$$

The space V^h consists of vector-valued quadratic polynomials with respect to the fine mesh such that there are two degrees of freedom on each vertex and each edge of this mesh. Considering the coarse mesh, then each mesh cell has the degrees of freedom on the vertices and edges and in addition the degrees of freedom on the barycenter and the three edges connecting the vertices with the barycenter. Altogether, one finds

$$\begin{aligned}\dim V^h &= 2\,(\#\,\text{vertices}_{\text{fine}} + \#\,\text{edges}_{\text{fine}})\\ &= 2\,(\#\,\text{vertices}_{\text{coarse}} + \#\,\text{edges}_{\text{coarse}} + 4\#\,\text{cells}_{\text{coarse}})\,.\end{aligned} \tag{4.193}$$

Using now the already shown exactness of the map **curl**, the rank-nullity theorem, Theorem A.64, Euler's formula for the triangulation of a simple closed polygon, see Remark C.10, (4.191), (4.192), and (4.193) yields

$$\begin{aligned}\dim\left(\text{div}\ V^h\right) &= \dim V^h - \dim\left(\ker\left(\text{div}\ V^h\right)\right) = \dim V^h - \dim\left(\textbf{curl}\ W^h\right)\\ &= \dim V^h - \dim W^h + \dim\left(\ker\left(\textbf{curl}\ W^h\right)\right)\\ &= \dim V^h - \dim W^h + 1\\ &= -\#\,\text{vertices}_{\text{coarse}} + \#\,\text{edges}_{\text{coarse}} + 8\#\,\text{cells}_{\text{coarse}} + 1\\ &= -\#\,\text{vertices}_{\text{coarse}} + \#\,\text{edges}_{\text{coarse}} + 8\#\,\text{cells}_{\text{coarse}}\\ &\quad\ +\#\,\text{vertices}_{\text{coarse}} - \#\,\text{edges}_{\text{coarse}} + \#\,\text{cells}_{\text{coarse}}\\ &= 9\#\,\text{cells}_{\text{coarse}} = \dim Q^h.\end{aligned}$$

Hence, $\text{div} : V^h \to Q^h$ is a surjection.

Using a macro-element technique from Arnold and Qin (1992), one can show that the mapping div has a bounded right inverse independent of h. Hence, the discrete inf-sup condition (3.51) is uniformly satisfied. □

Remark 4.145 (Other Pairs of Spaces) A similar construction as presented in Example 4.144 can be started with other $H^2(\Omega)$ conforming finite element spaces. However, the arising pairs of spaces for the Stokes equations seem to be of little importance in practical simulations due to the high polynomial degree of the basis functions or to the use of degrees of freedom whose implementation is involved, like gradients. Some examples of such spaces can be found, e.g., in Falk and Neilan (2013), Guzmán and Neilan (2014). □

Remark 4.146 (The Three-Dimensional Case) The situation in three dimensions is considerably more complicated than in two dimensions. Several Stokes complexes have been proposed. Due to the needed regularity of the spaces, the polynomial degree of the corresponding finite element spaces becomes very high, e.g., see Neilan (2015) where a velocity space with polynomials of degree six was derived.

□

Remark 4.147 ($H(\text{div}, \Omega)$ Conforming Finite Element Spaces with Weakly Divergence-free Velocity Solutions) Studying sub-complexes of the smooth de Rahm complex (4.189) leads to conforming finite element spaces since $V^h \subset H^1(\Omega)$ and $Q^h \subset L^2(\Omega)$. It turns out that one does not find lower order pairs of conforming spaces that are easily to implement and that can be used to compute weakly divergence-free velocity solutions.

In order to construct low order spaces, one has to abandon the conformity of the finite element velocity space, i.e., it has to hold $V^h \not\subset H^1(\Omega)$. Since one likes to have $\left\|\nabla \cdot \boldsymbol{u}^h\right\|_{L^2(\Omega)} = 0$, the divergence of the velocity fields has to be in $L^2(\Omega)$. This property is given if $V^h \subset H(\text{div}, \Omega)$. From Lemma 3.66, it is known that then the normal velocity is continuous across the faces of the mesh cells.

The perhaps simplest choice of finite element spaces in $H(\text{div}, \Omega)$ on simplicial meshes is $V^h = \text{RT}_0$, the Raviart–Thomas spaces of lowest order, see Example B.45. In fact, it can be shown that RT_0/P_0 satisfies a discrete inf-sup condition uniformly with respect to h where an appropriate norm of the functions from $V^h = \text{RT}_0$ is used, e.g., see (Boffi et al. 2013, Sect. 7.1.2). However, the non-conformity of V^h requires the definition of an appropriate discrete viscous term. The first idea consists in applying a piecewise definition as for the non-conforming Crouzeix–Raviart finite element spaces, see (4.18). However, it turns out that with this piecewise definition the consistency error for RT_0/P_0 is too large and the discrete solutions do not converge to the solution of the continuous problem. Thus, further modifications are needed to construct a convergent method.

In the literature, one finds two approaches for such modifications. Both approaches aim to reduce the order of the consistency error to the order of the best approximation error.

The first approach extends the piecewise defined bilinear form such that

- a term is introduced that accounts for the proper reduction of the consistency error,
- a second term establishes the symmetry of the discrete viscous term,
- a third term is responsible for a weak continuity of the tangential velocity at faces by penalizing jumps of the tangential velocity.

The second approach enriches the velocity space with local and weakly divergence-free functions in such a way that the tangential velocity becomes in a weak sense continuous.

More details of both approaches and references to the relevant literature can be found in the review paper (John et al. 2015, Sect. 4.4). □

Chapter 5
The Oseen Equations

Remark 5.1 (Motivation) Oseen equations, which are linear equations, show up as an auxiliary problem in many numerical approaches for solving the Navier–Stokes equations. Applying an implicit method for the temporal discretization of the Navier–Stokes equations requires the solution of a nonlinear problem in each discrete time. Likewise, the steady-state Navier–Stokes equations are nonlinear. Applying in either situation a so-called Picard method (a fixed point iteration) for solving the nonlinear problem, leads to an Oseen problem in each iteration, compare Sect. 6.3. The application of semi-implicit time discretizations to the Navier–Stokes equations leads directly to an Oseen problem in each discrete time, see Remark 7.61. Altogether, Oseen problems have to be solved in many methods for simulating the Navier–Stokes equations. In addition, some parts of the theory of the Oseen equations are used in the analysis of the Navier–Stokes equations, e.g., for the uniqueness of a weak solution of the steady-state Navier–Stokes equations in Theorem 6.20. For these reasons, the analysis and numerical analysis of Oseen problems is of fundamental interest.

In addition to the Stokes equations, the Oseen equations possess a convective term and a reactive term. Both of them might be dominant. Thus, besides the inf-sup condition, the third difficulty mentioned in Remark 2.19 has to be addressed.

□

5.1 The Continuous Equations

Remark 5.2 (The Oseen Equations) The Oseen equations are linear and stationary equations. In comparison with the Stokes equations (4.1), they have a convective term (first order derivative of the velocity) and they might possess also a so-called reactive term (zeroth order derivative of the velocity) in the momentum equation.

V. John, *Finite Element Methods for Incompressible Flow Problems*, Springer Series in Computational Mathematics 51, DOI 10.1007/978-3-319-45750-5_5

The Oseen equations have the form

$$\begin{aligned} -\nu\Delta \boldsymbol{u} + (\boldsymbol{b}\cdot\nabla)\boldsymbol{u} + c\boldsymbol{u} + \nabla p &= \boldsymbol{f} \text{ in } \Omega, \\ \nabla\cdot\boldsymbol{u} &= 0 \text{ in } \Omega, \end{aligned} \tag{5.1}$$

where $\nu > 0$, $\boldsymbol{b}$ is a given convection field which has to be in some sense divergence-free, see Remark 5.6 for details, $c(\boldsymbol{x}) \ge c_0 \ge 0$ for almost all $\boldsymbol{x} \in \Omega$, and $\boldsymbol{f} \in H^{-1}(\Omega)$.

For the application of the abstract theory from Chap. 3 and for the finite element error analysis, the Oseen equations will be equipped for simplicity with homogeneous Dirichlet boundary conditions $\boldsymbol{u} = \boldsymbol{0}$ on Γ. □

Remark 5.3 (On the Coefficients in (5.1)*)*

- If the functions in (5.1) are considered with dimensions, then all terms in the momentum equation have the unit $^{\mathrm{m}}/_{\mathrm{s}^2}$. Hence, the function c must have the unit $^1/_{\mathrm{s}}$, i.e., if $c(t,\boldsymbol{x}) > 0$, then $c^{-1}(t,\boldsymbol{x})$ represents a time, see also the discussion below in this remark.
- Let $\boldsymbol{b} \in L^\infty(\Omega)$, then it will be always assumed in this chapter that the momentum equation is scaled such that one of the following situations is given:
  - $\|\boldsymbol{b}\|_{L^\infty(\Omega)} \sim 1$ if $\nu \le \|\boldsymbol{b}\|_{L^\infty(\Omega)}$,
  - $\nu \sim 1$ if $\|\boldsymbol{b}\|_{L^\infty(\Omega)} \le \nu$.

  The first situation is the interesting one which appears in applications.
- The following cases for the other coefficients of (5.1) are of interest:
  - *ν is of moderate size, $c = 0$.* The case $c = 0$ occurs in numerical methods for solving the steady-state Navier–Stokes equations, e.g., if a fixed point iteration is applied, see Remark 6.41. From the point of view of applications, the limit $\nu \to 0$ is not of interest for the steady-state Navier–Stokes equations since for very small ν, or very large Reynolds numbers, a time-dependent solution is expected.

    However, strictly speaking, for finite element discretizations of the steady-state Navier–Stokes equations, the assumptions on the coefficients of the Oseen problem made in this chapter are usually not satisfied, e.g., see Remark 6.41.
  - *ν is of arbitrary size, $c = \mathcal{O}\left((\Delta t)^{-1}\right)$.* For small ν, one expects a time-dependent solution, for very small ν even a turbulent solution. In time-dependent problems, a term of type $c\boldsymbol{u}$ comes from an implicit or semi-implicit temporal discretization. Then, $c = \mathcal{O}\left((\Delta t)^{-1}\right)$, where Δt is the length of the time step, e.g., see (7.85). Thus, for small time steps, $c\boldsymbol{u}$ might become a dominant term in (5.1). Altogether, this case is of interest for a wide range of ν and a wide range of c.
- In numerical methods where Oseen problems appear, $\boldsymbol{b}$ is a computed velocity field, often the currently available finite element approximation of the solution.

Since generally $V^h_{\text{div}} \not\subset V_{\text{div}}$, this finite element approximation of the velocity is in general not weakly divergence-free. It is discretely divergence-free if an inf-sup stable pair of finite element spaces was used and if the algebraic saddle point problem was solved exactly. But even an exact solving is often not performed such that one can expect that the computed velocity field for $\boldsymbol{b}$ is only approximately discretely divergence-free.

□

Remark 5.4 (The Weak Form of the Oseen Equations) For the weak formulation of the Oseen problem (5.1) with homogeneous Dirichlet boundary conditions, the same functions spaces are used as for the weak form of the Stokes problem. Let $V = H_0^1(\Omega)$ and $Q = L_0^2(\Omega)$. Then, the weak form of (5.1) reads as follows: Find $(\boldsymbol{u}, p) \in V \times Q$ such that

$$\begin{aligned} \nu(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) + ((\boldsymbol{b} \cdot \nabla)\boldsymbol{u} + c\boldsymbol{u}, \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, p) &= \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{V',V} \quad \forall\ \boldsymbol{v} \in V, \\ -(\nabla \cdot \boldsymbol{u}, q) &= 0 \qquad\qquad\quad \forall\ q \in Q. \end{aligned} \tag{5.2}$$

To cast (5.2) into the abstract framework of Sect. 3.1, the following bilinear forms are defined:

$$\begin{aligned} a &: V \times V \to \mathbb{R}, \quad a(\boldsymbol{u}, \boldsymbol{v}) = \nu(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) + ((\boldsymbol{b} \cdot \nabla)\boldsymbol{u} + c\boldsymbol{u}, \boldsymbol{v}), \\ b &: V \times Q \to \mathbb{R}, \quad b(\boldsymbol{v}, q) = -(\nabla \cdot \boldsymbol{v}, q). \end{aligned} \tag{5.3}$$

The right-hand side of the second equation of the abstract problem (3.4) is $r = 0$.

□

Remark 5.5 (Minimal Regularity of $\boldsymbol{b}$ *and* c *for* (5.2) *To Be Well-posed)* Consider first the two-dimensional case. From the Sobolev imbedding (A.21), it follows that $\boldsymbol{u}, \boldsymbol{v} \in L^q(\Omega)$ with $q \in [1, \infty)$. Hence, it is sufficient that $\boldsymbol{b} \in L^{2+\varepsilon_0}(\Omega)$ and $c \in L^{1+\varepsilon_1}(\Omega)$, $\varepsilon_0, \varepsilon_1 > 0$, such that all terms in (5.2) are well defined.

In three dimensions, the Sobolev imbedding (A.22) gives $\boldsymbol{u}, \boldsymbol{v} \in L^6(\Omega)$ such that $\boldsymbol{b} \in L^3(\Omega)$ and $c \in L^{3/2}(\Omega)$ have to be satisfied.

In addition to the regularity assumptions, the properties on $\boldsymbol{b}$ and c given in Remark 5.2 have to be fulfilled. □

Remark 5.6 (Skew-Symmetry of the Convective Term) The key property of the convective term which will be used in the analysis is

$$((\boldsymbol{b} \cdot \nabla)\boldsymbol{v}, \boldsymbol{v}) = 0 \quad \forall\ \boldsymbol{v} \in V, \tag{5.4}$$

which is also called skew-symmetry.

- This property is given, e.g., if $\boldsymbol{b}$ satisfies the regularity assumptions from Remark 5.5, $\nabla \cdot \boldsymbol{b} \in L^2(\Omega)$, and $\nabla \cdot \boldsymbol{b} = 0$ almost everywhere in Ω, since integration by parts and the application of the product rule yields

$$((\boldsymbol{b} \cdot \nabla)\boldsymbol{v}, \boldsymbol{v}) = -(\nabla \cdot \boldsymbol{b}, \boldsymbol{v}^T\boldsymbol{v}) - ((\boldsymbol{b} \cdot \nabla)\boldsymbol{v}, \boldsymbol{v}). \tag{5.5}$$

The integral on Γ vanishes because $\boldsymbol{v} = \boldsymbol{0}$ on Γ. Since $\nabla \cdot \boldsymbol{b} = 0$, one obtains (5.4).

- Property (5.4) also holds if $\boldsymbol{b}$ and $\nabla \cdot \boldsymbol{b}$ satisfy the same regularity assumptions as in the previous case, $\boldsymbol{b}$ is weakly divergence-free, and

$$\int_\Gamma (\boldsymbol{b} \cdot \boldsymbol{n})(\boldsymbol{s})\, ds = 0. \tag{5.6}$$

For all $\boldsymbol{v} \in V$, it follows from the Sobolev imbeddings (A.21) and (A.22) that $\boldsymbol{v} \in L^4(\Omega)$. That means

$$\int_\Omega \|\boldsymbol{v}\|_2^4 \, d\boldsymbol{x} = \int_\Omega \left(\boldsymbol{v}^T \boldsymbol{v}\right)^2 \, d\boldsymbol{x} < \infty$$

and therefore that $\boldsymbol{v}^T\boldsymbol{v} \in L^2(\Omega)$. Then, there is a constant C such that $\boldsymbol{v}^T\boldsymbol{v} + C \in Q$. If $\boldsymbol{b}$ is weakly divergence-free, it holds

$$\begin{aligned} 0 &= \left(\nabla \cdot \boldsymbol{b}, \boldsymbol{v}^T\boldsymbol{v} + C\right) = \left(\nabla \cdot \boldsymbol{b}, \boldsymbol{v}^T\boldsymbol{v}\right) + (\nabla \cdot \boldsymbol{b}, C) \\ &= \left(\nabla \cdot \boldsymbol{b}, \boldsymbol{v}^T\boldsymbol{v}\right) - C\int_\Gamma \boldsymbol{b}\cdot\boldsymbol{n}\, ds, \end{aligned} \tag{5.7}$$

where in the last step integration by parts was applied. If (5.6) holds, then (5.4) follows from (5.5) and (5.7). A special case of this situation is that $\boldsymbol{b} \in V_{\text{div}}$.

- Under the same conditions on $\boldsymbol{b}$ as in the previous cases and with the same arguments, one finds that

$$((\boldsymbol{b}\cdot\nabla)\boldsymbol{v}, \boldsymbol{w}) = -((\boldsymbol{b}\cdot\nabla)\boldsymbol{w}, \boldsymbol{v}) \quad \forall\ \boldsymbol{v}, \boldsymbol{w} \in V. \tag{5.8}$$

□

Theorem 5.7 (Existence and Uniqueness of a Weak Solution of the Oseen Equations) *Let Ω be a bounded domain in $\mathbb{R}^d$, $d \in \{2, 3\}$, with a Lipschitz continuous boundary Γ and let the conditions on the data of the Oseen problem from Remarks 5.5 and 5.6 be fulfilled. Then, there exists a unique solution $(\boldsymbol{u}, p) \in H_0^1(\Omega) \times L_0^2(\Omega)$ of (5.2).*

Proof One has to check the conditions on $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ which are stated in Lemma 3.19.

With respect to the bilinear form $b(\cdot,\cdot)$, one has the same situation as for the Stokes problem, see Theorem 4.6. Hence, Theorem 3.46 states the fulfillment of the inf-sup condition.

For the bilinear form $a(\cdot,\cdot)$, one has to prove that $a(\cdot,\cdot)$ is coercive in V_{div}. Using (5.4), it follows that

$$a(\boldsymbol{v}, \boldsymbol{v}) = \nu(\nabla\boldsymbol{v}, \nabla\boldsymbol{v}) + (c\boldsymbol{v}, \boldsymbol{v}) \geq \nu \|\boldsymbol{v}\|_V^2,$$

because of $c \geq 0$, thus $a(\cdot,\cdot)$ is even coercive in V. Hence, the coercivity of $a(\cdot,\cdot)$ is proved and the application of Lemma 3.19 gives the statement of the theorem. ■

Lemma 5.8 (Stability of the Solution) *Under the conditions of Theorem 5.7, the solution of the Oseen problem* (5.2) *depends continuously on the data of the problem*

$$\frac{\nu}{2}\|\nabla \boldsymbol{u}\|^2_{L^2(\Omega)} + \left\|c^{1/2}\boldsymbol{u}\right\|^2_{L^2(\Omega)} \leq \frac{1}{2\nu}\|\boldsymbol{f}\|^2_{H^{-1}(\Omega)} \tag{5.9}$$

and if additionally $\boldsymbol{b}, c \in L^\infty(\Omega)$, *then*

$$\|p\|_{L^2(\Omega)} \leq \frac{1}{\beta_{\mathrm{is}}}\left[\|\boldsymbol{f}\|_{H^{-1}(\Omega)} + \left(\nu^{1/2} + C_{\mathrm{PF}}\frac{\|\boldsymbol{b}\|_{L^\infty(\Omega)}}{\nu^{1/2}} + C_{\mathrm{PF}}\|c\|^{1/2}_{L^\infty(\Omega)}\right)\right.$$
$$\left.\times\left(\nu^{1/2}\|\nabla \boldsymbol{u}\|_{L^2(\Omega)} + \left\|c^{1/2}\boldsymbol{u}\right\|_{L^2(\Omega)}\right)\right]. \tag{5.10}$$

If $\boldsymbol{f} \in L^2(\Omega)$ *and* $c_0 > 0$, *then also the stability bound*

$$\nu\|\nabla \boldsymbol{u}\|^2_{L^2(\Omega)} + \frac{1}{2}\left\|c^{1/2}\boldsymbol{u}\right\|^2_{L^2(\Omega)} \leq \frac{1}{2c_0}\|\boldsymbol{f}\|^2_{L_2(\Omega)} \tag{5.11}$$

is satisfied. If in addition $\boldsymbol{b}, c \in L^\infty(\Omega)$, *then*

$$\|p\|_{L^2(\Omega)} \leq \frac{1}{\beta_{\mathrm{is}}}\left[C_{\mathrm{PF}}\|\boldsymbol{f}\|_{L^2(\Omega)} + \left(\nu^{1/2} + \frac{\|\boldsymbol{b}\|_{L^\infty(\Omega)}}{c_0^{1/2}} + C_{\mathrm{PF}}\|c\|^{1/2}_{L^\infty(\Omega)}\right)\right.$$
$$\left.\times\left(\nu^{1/2}\|\nabla \boldsymbol{u}\|_{L^2(\Omega)} + \left\|c^{1/2}\boldsymbol{u}\right\|_{L^2(\Omega)}\right)\right] \tag{5.12}$$

holds. Here, C_{PF} *is the constant of the Poincaré–Friedrichs inequality* (A.12).

Proof The proof proceeds along the same lines as the proof of Lemma 4.7. To highlight the dependency of the stability bounds on the coefficients of the problem, it will be presented in detail.

Inserting $\boldsymbol{u} \in V$ as test function in (5.2), using (5.4), applying the estimate for the dual pairing, and using Young's inequality (A.5) leads to

$$\begin{aligned}\nu\|\nabla \boldsymbol{u}\|^2_{L^2(\Omega)} + \left\|c^{1/2}\boldsymbol{u}\right\|^2_{L^2(\Omega)} &\leq \|\boldsymbol{f}\|_{H^{-1}(\Omega)}\|\nabla \boldsymbol{u}\|_{L^2(\Omega)} \\ &\leq \frac{1}{2\nu}\|\boldsymbol{f}\|^2_{H^{-1}(\Omega)} + \frac{\nu}{2}\|\nabla \boldsymbol{u}\|^2_{L^2(\Omega)}.\end{aligned}$$

This inequality gives (5.9).

If $\boldsymbol{f} \in L^2(\Omega)$ and $c(\boldsymbol{x}) \geq c_0 > 0$, then the velocity estimate can be performed by applying the Cauchy–Schwarz inequality (A.10)

$$\begin{aligned}\nu \|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^2 + \left\|c^{1/2}\boldsymbol{u}\right\|_{L^2(\Omega)}^2 &\leq \|\boldsymbol{f}\|_{L_2(\Omega)} \|\boldsymbol{u}\|_{L^2(\Omega)} \leq \|\boldsymbol{f}\|_{L_2(\Omega)} \left\|\left(\frac{c}{c_0}\right)^{1/2} \boldsymbol{u}\right\|_{L^2(\Omega)} \\ &= \frac{1}{c_0^{1/2}} \|\boldsymbol{f}\|_{L_2(\Omega)} \left\|c^{1/2}\boldsymbol{u}\right\|_{L^2(\Omega)} \\ &\leq \frac{1}{2c_0} \|\boldsymbol{f}\|_{L_2(\Omega)}^2 + \frac{1}{2} \left\|c^{1/2}\boldsymbol{u}\right\|_{L^2(\Omega)}^2 .\end{aligned}$$

Now (5.11) follows.

The stability estimate for the pressure starts with the inf-sup condition. Then, Eq. (5.2) is inserted, (5.8) is applied, the estimate for the dual pairing, the Cauchy–Schwarz inequality, and the Poincaré–Friedrichs inequality are applied to obtain

$$\begin{aligned}\|p\|_{L^2(\Omega)} &\leq \frac{1}{\beta_{\text{is}}} \sup_{\boldsymbol{v}\in V\setminus\{\boldsymbol{0}\}} \frac{-(\nabla\cdot\boldsymbol{v}, p)}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)}} \\ &= \frac{1}{\beta_{\text{is}}} \sup_{\boldsymbol{v}\in V\setminus\{\boldsymbol{0}\}} \frac{\langle \boldsymbol{f}, \boldsymbol{v}\rangle_{V',V} - \nu(\nabla\boldsymbol{u}, \nabla\boldsymbol{v}) + ((\boldsymbol{b}\cdot\nabla)\boldsymbol{v}, \boldsymbol{u}) - (c\boldsymbol{u}, \boldsymbol{v})}{\|\nabla \boldsymbol{v}\|_{L^2(\Omega)}} \\ &\leq \frac{1}{\beta_{\text{is}}} \Bigg(\|\boldsymbol{f}\|_{H^{-1}(\Omega)} + \nu \|\nabla\boldsymbol{u}\|_{L^2(\Omega)} + \|\boldsymbol{b}\|_{L^\infty(\Omega)} \|\boldsymbol{u}\|_{L^2(\Omega)} \\ &\qquad + C_{\text{PF}} \|c\|_{L^\infty(\Omega)}^{1/2} \left\|c^{1/2}\boldsymbol{u}\right\|_{L^2(\Omega)} \Bigg).\end{aligned}$$

In order to be able to apply now the stability estimate for the velocity, $\|\boldsymbol{u}\|_{L^2(\Omega)}$ is estimated with the Poincaré–Friedrichs inequality, because the term $\|\boldsymbol{u}\|_{L^2(\Omega)}$ cannot be estimated with $\left\|c^{1/2}\boldsymbol{u}\right\|_{L^2(\Omega)}$ without the uniform positivity of $c(\boldsymbol{x})$. After having applied the Poincaré–Friedrichs inequality, the bound (5.9) is inserted, which gives (5.10).

In the case $\boldsymbol{f} \in L^2(\Omega)$ and $c(\boldsymbol{x}) \geq c_0 > 0$, one gets in the same way the estimate

$$\begin{aligned}\|p\|_{L^2(\Omega)} &\leq \frac{1}{\beta_{\text{is}}} \Bigg(C_{\text{PF}} \|\boldsymbol{f}\|_{L^2(\Omega)} + \nu \|\nabla\boldsymbol{u}\|_{L^2(\Omega)} + \frac{\|\boldsymbol{b}\|_{L^\infty(\Omega)}}{c_0^{1/2}} \left\|c^{1/2}\boldsymbol{u}\right\|_{L^2(\Omega)} \\ &\qquad + C_{\text{PF}} \|c\|_{L^\infty(\Omega)}^{1/2} \left\|c^{1/2}\boldsymbol{u}\right\|_{L^2(\Omega)} \Bigg),\end{aligned}$$

from which (5.12) follows by inserting (5.11). ■

Remark 5.9 (Discussion of the Stability Estimates) For the linearization of the steady-state Navier–Stokes equations with a fixed point iteration, where $c = 0$, estimates (5.9) and (5.10) are of interest. In particular, one has to study the dependency of the stability bounds on the viscosity. It can be seen that the bounds for $\|\nabla \boldsymbol{u}\|_{L^2(\Omega)}$ and $\|p\|_{L^2(\Omega)}$ scale like $\mathcal{O}\left(\nu^{-1}\right)$. Even replacing $\|\boldsymbol{f}\|_{H^{-1}(\Omega)}$ by $\|\boldsymbol{f}\|_{L^2(\Omega)}$, if this regularity of the right-hand side is given, does not change the dependency of the stability bounds on ν. Thus, for small ν, the problem loses stability. But note that the limit $\nu \to 0$ is not of that much interest in applications, see Remark 5.3.

For numerical schemes applied to the time-dependent Navier–Stokes equations, where the regularity $\boldsymbol{f} \in L^2(\Omega)$ is often given in applications, estimates (5.11) and (5.12) are relevant for $c = c_0 = \mathcal{O}\left((\Delta t)^{-1}\right)$. In this case, the stability bound for $\|\nabla \boldsymbol{u}\|_{L^2(\Omega)}$ scales like $\mathcal{O}\left(\left(\frac{\Delta t}{\nu}\right)^{1/2}\right)$, for $\|\boldsymbol{u}\|_{L^2(\Omega)}$ like $\mathcal{O}\left(\Delta t\right)$, and for $\|p\|_{L^2(\Omega)}$ like

$$\mathcal{O}\left(\left(1 + \left(\nu^{1/2} + (\Delta t)^{1/2} + (\Delta t)^{-1/2}\right)\left(\nu^{1/2}\left(\frac{\Delta t}{\nu}\right)^{1/2} + (\Delta t)^{-1/2}(\Delta t)\right)\right)\right)$$
$$= \mathcal{O}\left((\nu \Delta t)^{1/2} + \Delta t + 1\right).$$

If the length of the time step is sufficiently small, i.e., $\Delta t \leq \nu \leq 1$, then the stability bound for $\|\nabla \boldsymbol{u}\|_{L^2(\Omega)}$ behaves like $\mathcal{O}\left(1\right)$. For small time steps, also the two other stability bounds are small. □

5.2 The Galerkin Finite Element Method

Remark 5.10 (Goals) Besides existence, uniqueness, and stability of a solution of the finite element problem, finite element error estimates are the main topic of this section. The goals and principal approach for deriving such estimates are described in Remark 4.15. For the Oseen problem, the constants in the finite element error estimates will depend in particular on the coefficients of the equations. This dependency has to be tracked during performing the estimates. □

Remark 5.11 (The Galerkin Finite Element Method for Conforming Inf-Sup Stable Pairs of Finite Element Spaces) Let $V^h \subset V$, $Q^h \subset Q$ be inf-sup stable finite element spaces. Then, the Galerkin finite element method of (5.2) reads as follows: Find $\left(\boldsymbol{u}^h, p^h\right) \in V^h \times Q^h$ such that

$$\begin{aligned} a\left(\boldsymbol{u}^h, \boldsymbol{v}^h\right) + b\left(\boldsymbol{v}^h, p^h\right) &= \langle \boldsymbol{f}, \boldsymbol{v}^h\rangle_{V',V} \quad \forall\ \boldsymbol{v}^h \in V^h, \\ b\left(\boldsymbol{u}^h, q^h\right) &= 0 \qquad\qquad\quad \forall\ q^h \in Q^h, \end{aligned} \tag{5.13}$$

with the bilinear forms defined in (5.3). □

Corollary 5.12 (Unique Solvability of the Finite Element Problem) *Let the assumptions be as in Theorem 5.7, let Ω be a domain with polyhedral boundary, and let V^h and Q^h be conforming and inf-sup stable finite element spaces. Then, the finite element problem (5.13) has a unique solution.*

Proof The statement of the corollary follows from Lemma 3.19 by combining the coercivity of $a(\cdot,\cdot)$ in V^h, see the proof of Theorem 5.7, and the discrete inf-sup condition. ■

Lemma 5.13 (Stability of the Finite Element Solution) *Let $V^h \times Q^h$ be inf-sup stable finite element spaces. Then, the solution of (5.13) fulfills*

$$\frac{\nu}{2}\left\|\nabla \boldsymbol{u}^h\right\|_{L^2(\Omega)}^2 + \left\|c^{1/2}\boldsymbol{u}^h\right\|_{L^2(\Omega)}^2 \le \frac{1}{2\nu}\left\|\boldsymbol{f}\right\|_{H^{-1}(\Omega)}^2 \tag{5.14}$$

and if in addition $\boldsymbol{b}, c \in L^\infty(\Omega)$, then

$$\begin{aligned}\left\|p^h\right\|_{L^2(\Omega)} \le \frac{1}{\beta_{\text{is}}^h}\Bigg[&\left\|\boldsymbol{f}\right\|_{H^{-1}(\Omega)} + \left(\nu^{1/2} + C_{\text{PF}}\frac{\left\|\boldsymbol{b}\right\|_{L^\infty(\Omega)}}{\nu^{1/2}} + C_{\text{PF}}\left\|c\right\|_{L^\infty(\Omega)}^{1/2}\right)\\ &\times\left(\nu^{1/2}\left\|\nabla\boldsymbol{u}^h\right\|_{L^2(\Omega)} + \left\|c^{1/2}\boldsymbol{u}^h\right\|_{L^2(\Omega)}\right)\Bigg].\end{aligned}$$

If $\boldsymbol{f} \in L^2(\Omega)$ and $c(\boldsymbol{x}) \ge c_0 > 0$, then also the following stability estimates hold

$$\nu\left\|\nabla \boldsymbol{u}^h\right\|_{L^2(\Omega)}^2 + \frac{1}{2}\left\|c^{1/2}\boldsymbol{u}^h\right\|_{L^2(\Omega)}^2 \le \frac{1}{2c_0}\left\|\boldsymbol{f}\right\|_{L_2(\Omega)}^2$$

and, if $\boldsymbol{b}, c \in L^\infty(\Omega)$, then

$$\begin{aligned}\left\|p^h\right\|_{L^2(\Omega)} \le \frac{1}{\beta_{\text{is}}^h}\Bigg[&C_{\text{PF}}\left\|\boldsymbol{f}\right\|_{L^2(\Omega)} + \left(\nu^{1/2} + \frac{\left\|\boldsymbol{b}\right\|_{L^\infty(\Omega)}}{c_0^{1/2}} + C_{\text{PF}}\left\|c\right\|_{L^\infty(\Omega)}^{1/2}\right)\\ &\times\left(\nu^{1/2}\left\|\nabla\boldsymbol{u}^h\right\|_{L^2(\Omega)} + \left\|c^{1/2}\boldsymbol{u}^h\right\|_{L^2(\Omega)}\right)\Bigg].\end{aligned}$$

The constant C_{PF} comes from the Poincaré–Friedrichs inequality.

Proof The proof is performed exactly like the proof of Lemma 5.8. ■

Theorem 5.14 (Finite Element Error Estimate for the Velocity) *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, be a bounded domain with polyhedral and Lipschitz continuous boundary. Let $(\boldsymbol{u}, p) \in V \times Q$ be the unique solution of the Oseen problem (5.2) where the conditions on the data from Remarks 5.5 and 5.6 are assumed to be satisfied and let in addition $\boldsymbol{b}, c \in L^\infty(\Omega)$. Consider inf-sup stable conforming finite element spaces $V^h \times Q^h$ for discretizing (5.2) and denote by $\boldsymbol{u}^h \in V_{\text{div}}^h$ the velocity* *solution. Then, the following error estimate holds*

$$
\begin{aligned}
& \nu^{1/2} \left\| \nabla \left(\boldsymbol{u} - \boldsymbol{u}^h \right) \right\|_{L^2(\Omega)} + \left\| c^{1/2} \left(\boldsymbol{u} - \boldsymbol{u}^h \right) \right\|_{L^2(\Omega)} \\
& \le C \left[C_{\mathrm{os}} \inf_{\boldsymbol{v}^h \in V^h} \left\| \nabla (\boldsymbol{u} - \boldsymbol{v}^h) \right\|_{L^2(\Omega)} + \frac{1}{\nu^{1/2}} \inf_{q^h \in Q^h} \left\| p - q^h \right\|_{L^2(\Omega)} \right],
\end{aligned} \tag{5.15}
$$

where

$$
C_{\mathrm{os}} = \nu^{1/2} + \|c\|_{L^\infty(\Omega)}^{1/2} + \|\boldsymbol{b}\|_{L^\infty(\Omega)} \min \left\{ \frac{1}{\nu^{1/2}}, \frac{1}{c_0^{1/2}} \right\} \tag{5.16}
$$

and C does not depend on the coefficients of the Oseen problem.

Proof The principal ideas for proving the error estimate are the same as in the proof of the corresponding estimate for the Stokes problem, see Theorem 4.21. For the Oseen problem, there are some additional terms to be estimated and it is important to study the dependency of the error bound on the coefficients of the equations.

From the abstract theory of linear saddle point problems, Remark 3.11, it is known that the discrete velocity can be computed by solving the problem: Find $\boldsymbol{u}^h \in V_{\mathrm{div}}^h$ such that

$$
\nu \left(\nabla \boldsymbol{u}^h, \nabla \boldsymbol{v}^h \right) + \left((\boldsymbol{b} \cdot \nabla) \boldsymbol{u}^h + c \boldsymbol{u}^h, \boldsymbol{v}^h \right) = \langle \boldsymbol{f}, \boldsymbol{v}^h \rangle_{V',V} \ \forall \ \boldsymbol{v}^h \in V_{\mathrm{div}}^h. \tag{5.17}
$$

Using the functions from $V_{\mathrm{div}}^h \subset V$ as test functions in the continuous equation (5.2) and subtracting (5.17) gives the error equation

$$
\begin{aligned}
& \nu \left(\nabla (\boldsymbol{u} - \boldsymbol{u}^h), \nabla \boldsymbol{v}^h \right) + \left((\boldsymbol{b} \cdot \nabla) \left(\boldsymbol{u} - \boldsymbol{u}^h \right) + c \left(\boldsymbol{u} - \boldsymbol{u}^h \right), \boldsymbol{v}^h \right) \\
& - \left(\nabla \cdot \boldsymbol{v}^h, p - q^h \right) = 0 \quad \forall \ \boldsymbol{v}^h \in V_{\mathrm{div}}^h, \ \forall \ q^h \in Q^h.
\end{aligned} \tag{5.18}
$$

The error is split into a best approximation error in V_{div}^h and the difference of the best approximation to the solution of (5.17)

$$
\boldsymbol{u} - \boldsymbol{u}^h = \left(\boldsymbol{u} - I^h \boldsymbol{u} \right) - \left(\boldsymbol{u}^h - I^h \boldsymbol{u} \right) = \boldsymbol{\eta} - \boldsymbol{\phi}^h, \quad I^h \boldsymbol{u} \in V_{\mathrm{div}}^h.
$$

Now, $\boldsymbol{\phi}^h \in V_{\mathrm{div}}^h$ is used as test function in (5.18). Rearranging terms and using (5.4) yields

$$
\begin{aligned}
& \nu \left\| \nabla \boldsymbol{\phi}^h \right\|_{L^2(\Omega)}^2 + \left\| c^{1/2} \boldsymbol{\phi}^h \right\|_{L^2(\Omega)}^2 \\
& = \nu \left(\nabla \boldsymbol{\eta}, \nabla \boldsymbol{\phi}^h \right) + \left((\boldsymbol{b} \cdot \nabla) \boldsymbol{\eta} + c \boldsymbol{\eta}, \boldsymbol{\phi}^h \right) - \left(\nabla \cdot \boldsymbol{\phi}^h, p - q^h \right).
\end{aligned} \tag{5.19}
$$

The first term on the right-hand side of (5.19) is estimated by the Cauchy–Schwarz inequality (A.10) and Young's inequality (A.5)

$$\nu \left|\left(\nabla \boldsymbol{\eta}, \nabla \boldsymbol{\phi}^h\right)\right| \leq \nu \left\|\nabla \boldsymbol{\eta}\right\|_{L^2(\Omega)} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \leq \nu \left\|\nabla \boldsymbol{\eta}\right\|_{L^2(\Omega)}^2 + \frac{\nu}{4} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 .$$

Similarly, one obtains for the last term with (3.41)

$$\left|-\left(\nabla \cdot \boldsymbol{\phi}^h, p - q^h\right)\right| \leq \frac{2}{\nu} \left\|p - q^h\right\|_{L^2(\Omega)}^2 + \frac{\nu}{8} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 ,$$

where in addition (3.40) was used. Also the reactive term is estimated in this way

$$\begin{aligned} \left|\left(c\boldsymbol{\eta}, \boldsymbol{\phi}^h\right)\right| &\leq \left\|c\right\|_{L^\infty(\Omega)}^{1/2} \left\|\boldsymbol{\eta}\right\|_{L^2(\Omega)} \left\|c^{1/2}\boldsymbol{\phi}\right\|_{L^2(\Omega)} \\ &\leq \left\|c\right\|_{L^\infty(\Omega)} \left\|\boldsymbol{\eta}\right\|_{L^2(\Omega)}^2 + \frac{1}{4} \left\|c^{1/2}\boldsymbol{\phi}\right\|_{L^2(\Omega)}^2 . \end{aligned}$$

For the convective term, one obtains with (5.8)

$$\begin{aligned} \left|\left(\left(\boldsymbol{b} \cdot \nabla\right) \boldsymbol{\eta}, \boldsymbol{\phi}^h\right)\right| = \left|-\left(\left(\boldsymbol{b} \cdot \nabla\right) \boldsymbol{\phi}^h, \boldsymbol{\eta}\right)\right| &\leq \left\|\boldsymbol{b}\right\|_{L^\infty(\Omega)} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \left\|\boldsymbol{\eta}\right\|_{L^2(\Omega)} \\ &\leq \frac{2}{\nu} \left\|\boldsymbol{b}\right\|_{L^\infty(\Omega)}^2 \left\|\boldsymbol{\eta}\right\|_{L^2(\Omega)}^2 + \frac{\nu}{8} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 . \end{aligned}$$

In the case $c_0 > 0$, there is the alternative estimate

$$\begin{aligned} \left|\left(\left(\boldsymbol{b} \cdot \nabla\right) \boldsymbol{\eta}, \boldsymbol{\phi}^h\right)\right| &\leq \left\|\boldsymbol{b}\right\|_{L^\infty(\Omega)} \left\|c^{-1/2}\right\|_{L^\infty(\Omega)} \left\|\nabla \boldsymbol{\eta}\right\|_{L^2(\Omega)} \left\|c^{1/2}\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \\ &\leq \frac{\left\|\boldsymbol{b}\right\|_{L^\infty(\Omega)}^2 \left\|\nabla \boldsymbol{\eta}\right\|_{L^2(\Omega)}^2}{c_0} + \frac{\left\|c^{1/2}\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2}{4}. \end{aligned}$$

Inserting all estimates in (5.19) gives

$$\begin{aligned} &\frac{1}{2} \left(\nu \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 + \left\|c^{1/2}\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2\right) \\ &\leq \nu \left\|\nabla \boldsymbol{\eta}\right\|_{L^2(\Omega)}^2 + \left\|c\right\|_{L^\infty(\Omega)} \left\|\boldsymbol{\eta}\right\|_{L^2(\Omega)}^2 \\ &\quad + \min\left\{\frac{2 \left\|\boldsymbol{b}\right\|_{L^\infty(\Omega)}^2 \left\|\boldsymbol{\eta}\right\|_{L^2(\Omega)}^2}{\nu}, \frac{\left\|\boldsymbol{b}\right\|_{L^\infty(\Omega)}^2 \left\|\nabla \boldsymbol{\eta}\right\|_{L^2(\Omega)}^2}{c_0}\right\} + \frac{2}{\nu} \left\|p - q^h\right\|_{L^2(\Omega)}^2 . \end{aligned} \tag{5.20}$$

From the triangle inequality, inequality (A.4), and the Poincaré–Friedrichs inequality (A.12), one obtains

$$\begin{aligned}
&\nu^{1/2} \left\| \nabla \left(\boldsymbol{u} - \boldsymbol{u}^h \right) \right\|_{L^2(\Omega)} + \left\| c^{1/2} \left(\boldsymbol{u} - \boldsymbol{u}^h \right) \right\|_{L^2(\Omega)} \\
&\le \nu^{1/2} \left\| \nabla \boldsymbol{\phi}^h \right\|_{L^2(\Omega)} + \left\| c^{1/2} \boldsymbol{\phi}^h \right\|_{L^2(\Omega)} + \nu^{1/2} \left\| \nabla \boldsymbol{\eta} \right\|_{L^2(\Omega)} + \left\| c^{1/2} \boldsymbol{\eta} \right\|_{L^2(\Omega)} \\
&\le \sqrt{2} \left(\nu \left\| \nabla \boldsymbol{\phi}^h \right\|_{L^2(\Omega)}^2 + \left\| c^{1/2} \boldsymbol{\phi}^h \right\|_{L^2(\Omega)}^2 \right)^{1/2} + \nu^{1/2} \left\| \nabla \boldsymbol{\eta} \right\|_{L^2(\Omega)} + \left\| c \right\|_{L^\infty(\Omega)}^{1/2} \left\| \boldsymbol{\eta} \right\|_{L^2(\Omega)} \\
&\le \sqrt{2} \left(\nu \left\| \nabla \boldsymbol{\phi}^h \right\|_{L^2(\Omega)}^2 + \left\| c^{1/2} \boldsymbol{\phi}^h \right\|_{L^2(\Omega)}^2 \right)^{1/2} + \left(\nu^{1/2} + C_{\mathrm{PF}} \left\| c \right\|_{L^\infty(\Omega)}^{1/2} \right) \left\| \nabla \boldsymbol{\eta} \right\|_{L^2(\Omega)} .
\end{aligned}$$

Inserting (5.20), applying once more (A.4) the Poincaré–Friedrichs inequality, and using (3.71) if possible, or (3.65) otherwise, gives (5.15). ■

Theorem 5.15 (Finite Element Error Estimate for the $L^2(\Omega)$ Norm of the Pressure) *Let the assumption of Theorem 5.14 hold. Then the following error estimate holds for the pressure*

$$\begin{aligned}
\left\| p - p^h \right\|_{L^2(\Omega)} \le C \Bigg[& \frac{1}{\beta_{\mathrm{is}}^h} C_{\mathrm{os}}^2 \inf_{\boldsymbol{v}^h \in V^h} \left\| \nabla (\boldsymbol{u} - \boldsymbol{v}^h) \right\|_{L^2(\Omega)} \\
& + \left(1 + \frac{1}{\beta_{\mathrm{is}}^h} + \frac{1}{\beta_{\mathrm{is}}^h} \frac{C_{\mathrm{os}}}{\nu^{1/2}} \right) \inf_{q^h \in Q^h} \left\| p - q^h \right\|_{L^2(\Omega)} \Bigg],
\end{aligned} \tag{5.21}$$

where C_{os} is defined in (5.16) and C does not depend on the coefficients of the Oseen equations.

Proof The way of proving (5.21) follows the proof of Theorem 4.25.

Using the discrete inf-sup condition (3.51), the discrete Oseen equations (5.13) as well as the continuous Oseen equations (5.2), one obtains in the same way as in the proof of Theorem 4.25

$$\left\| p^h - q^h \right\|_{L^2(\Omega)} \le \frac{1}{\beta_{\mathrm{is}}^h} \sup_{\boldsymbol{v}^h \in V^h \setminus \{\boldsymbol{0}\}} \frac{a\left(\boldsymbol{u} - \boldsymbol{u}^h, \boldsymbol{v}^h \right) + b\left(\boldsymbol{v}^h, p - q^h \right)}{\left\| \nabla \boldsymbol{v}^h \right\|_{L^2(\Omega)}} \tag{5.22}$$

for all $q^h \in Q^h$. The bilinear forms are replaced by (5.3) and then the individual terms are estimated. Using the Cauchy–Schwarz inequality (A.10), (3.41), and the Poincaré–Friedrichs inequality (A.12) yields

$$\begin{aligned}
\left| \nu \left(\nabla \left(\boldsymbol{u} - \boldsymbol{u}^h \right), \nabla \boldsymbol{v}^h \right) \right| &\le \nu \left\| \nabla \left(\boldsymbol{u} - \boldsymbol{u}^h \right) \right\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{v}^h \right\|_{L^2(\Omega)} , \\
\left| \left(\nabla \cdot \boldsymbol{v}^h, p - q^h \right) \right| &\le \left\| p - q^h \right\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{v}^h \right\|_{L^2(\Omega)} , \\
\left| \left(c \left(\boldsymbol{u} - \boldsymbol{u}^h \right), \boldsymbol{v}^h \right) \right| &\le C_{\mathrm{PF}} \left\| c \right\|_{L^\infty(\Omega)}^{1/2} \left\| c^{1/2} \left(\boldsymbol{u} - \boldsymbol{u}^h \right) \right\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{v}^h \right\|_{L^2(\Omega)} .
\end{aligned}$$

The convective term can be estimated with the same tools

$$\left|\left((\boldsymbol{b}\cdot\nabla)\left(\boldsymbol{u}-\boldsymbol{u}^h\right),\boldsymbol{v}^h\right)\right| \le C_{\mathrm{PF}}\,\|\boldsymbol{b}\|_{L^\infty(\Omega)}\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)}$$

or with applying integration by parts in the first step of the estimate, see (5.8),

$$\left|\left((\boldsymbol{b}\cdot\nabla)\left(\boldsymbol{u}-\boldsymbol{u}^h\right),\boldsymbol{v}^h\right)\right| \le \frac{\|\boldsymbol{b}\|_{L^\infty(\Omega)}}{c_0^{1/2}}\left\|c^{1/2}\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)}.$$

Inserting these estimates in (5.22) leads for all $q^h \in Q^h$ to

$$\begin{aligned}&\left\|p^h-q^h\right\|_{L^2(\Omega)}\\ &\le \frac{C}{\beta_{\mathrm{is}}^h}\left(\nu^{1/2}+\|c\|_{L^\infty(\Omega)}^{1/2}+\|\boldsymbol{b}\|_{L^\infty(\Omega)}\min\left\{\frac{1}{\nu^{1/2}},\frac{1}{c_0^{1/2}}\right\}\right)\\ &\quad\times\left(\nu^{1/2}\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}+\left\|c^{1/2}\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}\right)+\frac{C}{\beta_{\mathrm{is}}^h}\left\|p-q^h\right\|_{L^2(\Omega)}.\end{aligned}$$

The triangle inequality and the insertion of (5.15) conclude the proof. ■

Corollary 5.16 (Finite Element Error Estimates for Conforming Pairs of Finite Element Spaces) *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, be a bounded domain with polyhedral and Lipschitz continuous boundary which is decomposed by a regular and quasi-uniform family of triangulations $\{\mathcal{T}^h\}$. Let $(\boldsymbol{u},p)$ be the solution of the Oseen equations (5.2) with $\boldsymbol{u} \in H^{k+1}(\Omega)\cap V$ and $p \in H^k(\Omega)\cap Q$. Then for the inf-sup stable pairs of finite element spaces*

- $P_k^{\mathrm{bubble}}/P_k$, $k=1$ *(MINI element),*
- P_k/P_{k-1}, Q_k/Q_{k-1}, $k \ge 2$ *(Taylor–Hood element),*
- $P_2^{\mathrm{bubble}}/P_1^{\mathrm{disc}}$, $P_3^{\mathrm{bubble}}/P_2^{\mathrm{disc}}$, $P_2^{\mathrm{BR}}/P_1^{\mathrm{disc}}$, $Q_k/P_{k-1}^{\mathrm{disc}}$, $k \ge 2$,

the following error estimates hold

$$\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)} \le \frac{C}{\nu^{1/2}}h^k\left(C_{\mathrm{os}}\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}+\frac{1}{\nu^{1/2}}\|p\|_{H^k(\Omega)}\right), \tag{5.23}$$

where the constant depends either on the inverse of the discrete inf-sup constant β_{is}^h or on the inverse of local inf-sup constants, compare Remark 4.29, and

$$\left\|p-p^h\right\|_{L^2(\Omega)} \le Ch^k\left(C_{\mathrm{os}}^2\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}+\left(1+\frac{C_{\mathrm{os}}}{\nu^{1/2}}\right)\|p\|_{H^k(\Omega)}\right), \tag{5.24}$$

where the constant depends on the inverse of the discrete inf-sup constant.

Proof The estimates follow directly from estimating the best approximation errors in (5.15) and (5.21) with interpolation errors which are known for the finite element spaces, see Theorem C.13. The dependency of the constant of estimate (5.23) follows from the application of (3.71) or (3.65) in the last step of the proof of Theorem 5.14. ■

Remark 5.17 (Discussion of the Dependency of the Error Bounds on the Coefficients of the Problem) This discussion considers in particular the two cases of interest that were described in Remark 5.3. Possible dependencies of $\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}$ and $\|p\|_{H^k(\Omega)}$ on the coefficients of the problem are not taken into account. The dependency of the error estimates on the coefficients of the problem will be discussed only for the interesting case $\nu \le 1$.

- Consider first the case $c = c_0 = 0$, which appears in the iterative solution of the steady-state Navier–Stokes equations with a fixed point iteration, see Remark 6.41. In this case there is $C_{\mathrm{os}} = \mathcal{O}\left(\nu^{-1/2}\right)$, see Fig. 5.1. It follows that the constants in the estimates (5.23) and (5.24) behave both like $\mathcal{O}\left(\nu^{-1}\right)$ such that the error bounds blow up for $\nu \to 0$. As already mentioned in Remark 5.3, the case $\nu \to 0$ is not of interest in applications. However, ν can be nevertheless small and in this case the constants in the estimates (5.23) and (5.24) might become large.
- The other interesting situation is $c = c_0 = \mathcal{O}\left((\Delta t)^{-1}\right)$. This case arises in numerical methods for the time-dependent Navier–Stokes equations and thus also the situation that ν is very small is of interest.
  - ∘ If the time step is sufficiently small, i.e., $\Delta t \le \nu^{-1}$ and $\Delta t \le (\Delta t)^{-1}$, then $C_{\mathrm{os}} = \mathcal{O}\left((\Delta t)^{-1/2}\right)$. It follows that the constant in (5.23) scales like $\mathcal{O}\left((\nu\Delta t)^{-1/2}\right) + \mathcal{O}\left((\nu^{-1})\right)$ and the constant in (5.24) like $\mathcal{O}\left((\Delta t)^{-1}\right) + \mathcal{O}\left((\nu\Delta t)^{-1/2}\right)$.

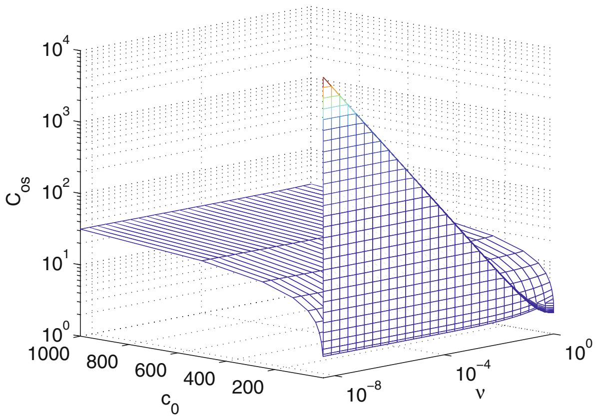

Fig. 5.1 The constant C_{os} from the error bounds (5.23) and (5.24) for $\|\boldsymbol{b}\|_{L^\infty(\Omega)} = 1$

- For large time steps compared with the viscosity, i.e., $\Delta t > \nu^{-1}$, one has $C_{\text{os}} = \mathcal{O}\left(\nu^{-1/2}\right)$, which leads to constants of size $\mathcal{O}\left(\nu^{-1}\right)$.

In all cases, the error bounds become large for a small viscosity ν (or a large Reynolds number). They become also large if $\|c\|_{L^\infty(\Omega)}$ is large, in particular for small time steps. In summary, the error bounds are not uniform with respect to the coefficients of the problem. Thus, it can happen that the Galerkin finite element discretization for the Oseen problem leads to large errors, compare Example 5.18. □

Example 5.18 (Analytical Example Which Supports the Error Estimates (5.23) *and* (5.24)*)* The numerical results that will be presented were obtained for Example D.3. That means, the solution and Sobolev norms of the solution do not depend on the viscosity. Of course, any divergence-free vector field can be chosen as convection field in the Oseen equations (5.1). Form the point of view of applications, the most important case is that the convection field is the velocity solution of the Oseen equations itself. This case will be considered here, i.e., the convection field is given by (D.9). Thus, $\|\boldsymbol{b}\|_{L^\infty(\Omega)} = \mathcal{O}(1)$.

The simulations were performed for the Taylor–Hood element Q_2/Q_1 on the quadrilateral grid depicted in Fig. 4.2. Results will be presented for the two situations of interest described in Remark 5.3.

Figure 5.2 presents the errors $\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}$ and $\left\|p-p^h\right\|_{L^2(\Omega)}$ for choosing $c = 0$ and different values of ν. The second order convergence for large values of the viscosity and the increase of the errors for small ν on coarse grids can be clearly seen. However, the order of error reduction for small ν on coarse grids is larger than two. On finer grids, the dependency of the error on ν becomes smaller. It can be expected that eventually, on sufficiently fine grids, the errors become independent of ν. With respect to the pressure, similar observations as for the velocity can be made.

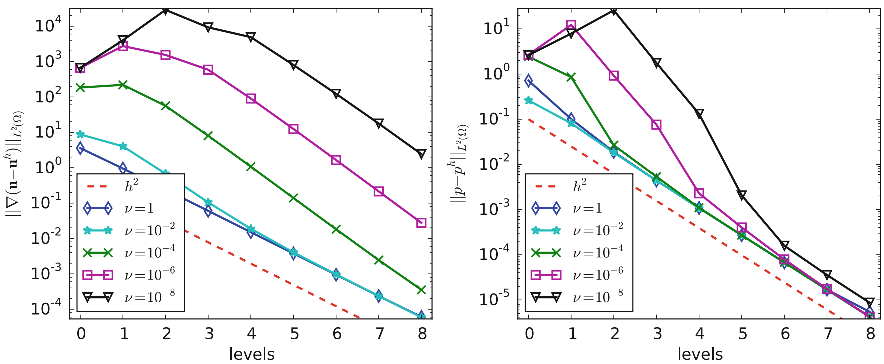

Fig. 5.2 Example 5.18. Convergence of the errors $\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)}$ and $\left\|p-p^h\right\|_{L^2(\Omega)}$ for $c = 0$ and different values of ν

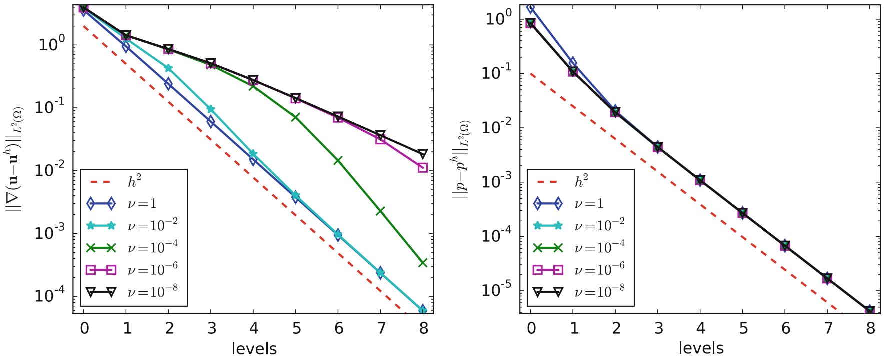

Fig. 5.3 Example 5.18. Convergence of the errors $\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\|_{L^2(\Omega)}$ and $\|p-p^h\|_{L^2(\Omega)}$ for $c = 100$ and different values of ν

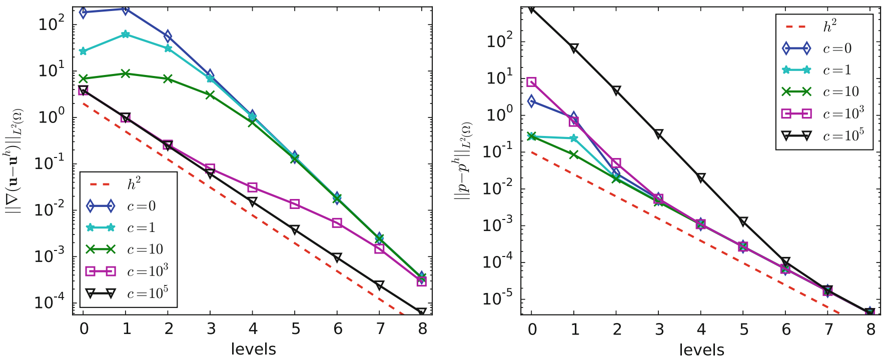

Fig. 5.4 Example 5.18. Convergence of the errors $\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\|_{L^2(\Omega)}$ and $\|p-p^h\|_{L^2(\Omega)}$ for $\nu = 10^{-4}$ and different values of c

In the second numerical study, the coefficient of the zeroth order term was chosen to be $c = 100$ and the viscosity was varied, see Fig. 5.3. With respect to the pressure, one obtains again more or less the same accuracy for all values of the viscosity. For the velocity, one gets larger errors for smaller values of ν. However, analogously to the case $c = 0$, it can be expected from the results for $\nu = 10^{-2}$ and $\nu = 10^{-4}$ that on sufficiently fine grids the errors become similar.

Finally, the case of $\nu = 10^{-4}$ and different values of c was studied, see Fig. 5.4. With respect to the error in the velocity, it can be seen that an increase of c (which corresponds to small time steps) decreases the error on coarse grids. On finer grids, the errors tend to become independent of c. Concerning the error in the pressure, one obtains almost the same results for all values of c on finer grids. On coarser grids, in the pre-asymptotic region, the largest errors can be observed for large values of c.

□

Remark 5.19 (Implementation) The principal approach concerning the construction of finite element spaces is the same as for the Stokes equations, see Sect. 4.3.

The linear saddle point problem for the Galerkin finite element discretization (5.13) has the form

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{0} \end{pmatrix}, \tag{5.25}$$

where the matrix block A is a diagonal block matrix with the same blocks

$$\begin{pmatrix} A_{11} & 0 & 0 \\ 0 & A_{11} & 0 \\ 0 & 0 & A_{11} \end{pmatrix}. \tag{5.26}$$

The entries are given by

$$(A_{11})_{ij} = a_{ij} = \sum_{K \in \mathcal{T}^h} \left[\left(\nu \nabla \boldsymbol{\phi}_j^h, \nabla \boldsymbol{\phi}_i^h\right)_K + \left((\boldsymbol{b} \cdot \nabla)\, \boldsymbol{\phi}_j^h + c\boldsymbol{\phi}_j^h, \boldsymbol{\phi}_i^h\right)_K \right],$$

$i, j = 1, \ldots, 3N_v$. Thus, if $\boldsymbol{b} \neq \boldsymbol{0}$, the matrix A is not symmetric. But the matrix structure is the same as for the Stokes problem, see Remark 4.66. □

5.3 Residual-Based Stabilizations

Remark 5.20 (Residual-Based Stabilizations) Residual-based methods are one of the most popular approaches for obtaining stable discretizations of the Oseen equations, or more general, of convection-dominated problems. Several types of residual-based stabilizations for the Oseen equations can be found in the literature, e.g., see Braack et al. (2007). In this section, a residual-based stabilization involving a Streamline-Upwind Petrov–Galerkin (SUPG) term, a pressure stabilization Petrov–Galerkin (PSPG) term, and a grad-div stabilization term, is presented in detail. □

5.3.1 The Basic Idea

Remark 5.21 (The Basic Idea of Residual-Based Stabilizations) The basic idea consists in a penalization of large values of the so-called strong residual.

Given a linear partial differential equation in strong form

$$A_{\mathrm{str}} u_{\mathrm{str}} = f, \quad f \in L^2(\Omega),$$

and its Galerkin finite element discretization: Find $u^h \in V^h$ such that

$$a^h\left(u^h, v^h\right) = \left(f, v^h\right) \quad \forall\ v^h \in V^h. \tag{5.27}$$

For residual-based stabilizations, a modification of A_{str} is needed which is well defined for finite element functions. This modification should be also a linear operator and it is denoted by $A^h_{\mathrm{str}} : V^h \to L^2(\Omega)$. The (strong) residual is now defined by

$$r^h\left(u^h\right) = A^h_{\mathrm{str}} u^h - f \ \in L^2(\Omega).$$

In general, it holds $r^h\left(u^h\right) \neq 0$, but a good numerical approximation of the solution of the continuous problem should have in some sense a small residual. Now, instead of finding the solution of (5.27), the minimizer of the residual is searched, i.e, the following optimization problem is considered

$$\arg\min_{u^h \in V^h} \left\| r^h\left(u^h\right)\right\|^2_{L^2(\Omega)} = \arg\min_{u^h \in V^h} \left(r^h\left(u^h\right), r^h\left(u^h\right)\right). \tag{5.28}$$

The necessary condition for taking the minimum is the vanishing of the Gâteaux derivative. This derivative is computed by using the linearity of A^h_{str} and the bilinearity of the inner product in $L^2(\Omega)$

$$\begin{aligned}
0 &= \lim_{\varepsilon \to 0} \frac{\left(r^h\left(u^h + \varepsilon v^h\right), r^h\left(u^h + \varepsilon v^h\right)\right) - \left(r^h\left(u^h\right), r^h\left(u^h\right)\right)}{\varepsilon} \\
&= \lim_{\varepsilon \to 0} \frac{\left(r^h\left(u^h\right) + \varepsilon A^h_{\mathrm{str}} v^h, r^h\left(u^h\right) + \varepsilon A^h_{\mathrm{str}} v^h\right) - \left(r^h\left(u^h\right), r^h\left(u^h\right)\right)}{\varepsilon} \\
&= 2\left(r^h\left(u^h\right), A^h_{\mathrm{str}} v^h\right) \quad \forall\ v^h \in V^h.
\end{aligned}$$

It follows that the necessary condition for the solution of (5.28) is

$$\left(r^h\left(u^h\right), A^h_{\mathrm{str}} v^h\right) = 0 \quad \forall\ v^h \in V^h.$$

A generalization consists in considering the minimization problem

$$\arg\min_{u^h \in V^h} \left\| \delta^{1/2} r^h\left(u^h\right)\right\|^2_{L^2(\Omega)} = \arg\min_{u^h \in V^h} \left(\delta r^h\left(u^h\right), r^h\left(u^h\right)\right) \tag{5.29}$$

with the positive weighting function $\delta(\boldsymbol{x})$. Analogously to the derivation for the special case, one obtains as necessary condition for the minimum

$$\left(\delta r^h\left(u^h\right), A^h_{\mathrm{str}} v^h\right) = 0 \quad \forall\ v^h \in V^h. \tag{5.30}$$

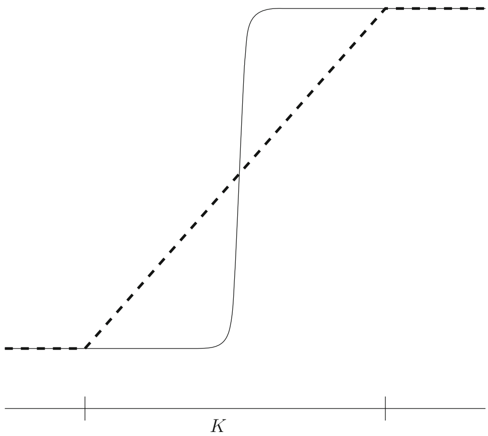

Fig. 5.5 Function with sharp layer (*solid line*) and optimal piecewise linear approximation in a mesh cell K (*dashed line*). The equation that is satisfied by the function in K is far from being satisfied by the piecewise linear approximation. Hence, despite the approximation is considered to be optimal, the strong residual will be large

The solutions of (5.28) or (5.29) will not be identical to the solution of the Galerkin discretization (5.27). It turns out that the reason for the Galerkin discretization to compute numerical solutions with large errors is that the solution of the continuous problem possesses structures (scales) that are important but that are not resolved by the used finite element space (grid), which is in particular the main difficulty in the simulation of turbulent flows, see Sect. 8.1. For flow problems, such structures are layers, particularly at boundaries. The numerical methods should compute sharp layers. However the sharpness of layers in numerical solutions is restricted by the resolution of the finite element spaces, which is generally much coarser than the layer width. Hence, even for a numerical solution with sharp layers, the strong residual in the layer regions is very large. In particular, a numerical solution with sharp layers (with respect to the resolution of the finite element space) will not be the minimizer of (5.28) or (5.29), see Fig. 5.5. The minimizer of (5.28) or (5.29) tends to possess strongly smeared layers and these solutions are often not useful in applications. For this reason, one considers in residual-based stabilizations a combination of the Galerkin discretization (5.27), which possesses too little numerical viscosity, and the minimization of the strong residual, which introduces too much numerical viscosity,

$$a^h\left(u^h, v^h\right) + \left(\delta r^h\left(u^h\right), A^h_{\mathrm{str}} v^h\right) = \left(f, v^h\right) \quad \forall\ v^h \in V^h. \tag{5.31}$$

The goal of the numerical analysis consists in determining the weighting function δ optimally. This goal is generally not achieved completely, only asymptotically optimal choices for weighting functions are known, e.g. see Remark 5.42. □

Example 5.22 (Oseen Equations) There are two equations in the Oseen equations and these equations should be treated separately. Given a triangulation $\mathcal{T}^h$ with mesh cells $\{K\}$. Then, one has for the momentum equation

$$A_{\text{str,m}}(\boldsymbol{u},p) = -\nu\Delta\boldsymbol{u} + (\boldsymbol{b}\cdot\nabla)\boldsymbol{u} + c\boldsymbol{u} + \nabla p,$$
$$A^h_{\text{str,m}}\left(\boldsymbol{u}^h,p^h\right) = -\nu\Delta\boldsymbol{u}^h + (\boldsymbol{b}\cdot\nabla)\boldsymbol{u}^h + c\boldsymbol{u}^h + \nabla p^h \quad \text{if} \quad \boldsymbol{x}\in\mathring{K}, \forall\, K\in\mathcal{T}^h.$$

Thus, $A^h_{\text{str,m}}$ is not defined on the faces of the mesh cells. However, the union of the faces is a set of Lebesgue measure zero and since $A^h_{\text{str,m}}\left(\boldsymbol{u}^h,p^h\right)\in L^2(\Omega)$, it is not necessary to define this function on this set. For the continuity equation, one has

$$A_{\text{str,c}}(\boldsymbol{u}) = -\nabla\cdot\boldsymbol{u},$$
$$A^h_{\text{str,c}}\left(\boldsymbol{u}^h\right) = -\nabla\cdot\boldsymbol{u}^h \quad \text{if} \quad \boldsymbol{x}\in\mathring{K}, \forall\, K\in\mathcal{T}^h.$$

The corresponding residuals are denoted by $r^h_m\left(\boldsymbol{u}^h,p^h\right)$ and $r^h_c\left(\boldsymbol{u}^h\right)$, respectively.

The general framework for minimizing the residual leads to the following terms in (5.31), besides the terms coming from the Galerkin discretization,

$$\begin{aligned}
&\sum_{K\in\mathcal{T}^h}\left[\left(\delta_c r^h_c\left(\boldsymbol{u}^h\right), A^h_{\text{str,c}}(\boldsymbol{v}^h)\right)_K + \left(\delta_m r^h_m\left(\boldsymbol{u}^h,p^h\right), A^h_{\text{str,m}}(\boldsymbol{v}^h,q^h)\right)_K\right]\\
&\quad = \sum_{K\in\mathcal{T}^h}\left[\left(\delta_c\nabla\cdot\boldsymbol{u}^h,\nabla\cdot\boldsymbol{v}^h\right)_K\right.\\
&\qquad +\Big(\delta_m\left(-\nu\Delta\boldsymbol{u}^h + (\boldsymbol{b}\cdot\nabla)\boldsymbol{u}^h + c\boldsymbol{u}^h + \nabla p^h - \boldsymbol{f}\right),\\
&\qquad \left.-\nu\Delta\boldsymbol{v}^h + (\boldsymbol{b}\cdot\nabla)\boldsymbol{v}^h + c\boldsymbol{v}^h + \nabla q^h\Big)_K\right].
\end{aligned} \tag{5.32}$$

The decomposition of the integrals is necessary since the terms are generally not well defined on the boundaries of the mesh cells. The first term on the right-hand side of (5.32) is the grad-div term, which was already studied in Sect. 4.6.1.

The terms in (5.32) are the prototype of residual-based stabilization terms for the Oseen equations. The actually used residual-based stabilizations contain some modifications. □

5.3.2 *The SUPG/PSPG/grad-div Stabilization*

Remark 5.23 (Plan of this Section) The numerical analysis of stabilized methods becomes technically much more involved than the numerical analysis of the Galerkin discretization. The contents of this section and the main steps of the numerical analysis are as follows:

- introduction of the stabilized method, Remark 5.24 and some general discussions of this method,

- consistency and Galerkin orthogonality of the method, Lemma 5.31,
- introduction of a norm for the numerical analysis, which is almost without pressure contribution, in (5.40),
- existence and uniqueness of a solution for this method, Theorem 5.34, and stability of the solution, Lemma 5.35,
- introduction of a norm with some pressure contribution in (5.48),
- proof of a discrete inf-sup condition for the bilinear form of the method where the constant does not depend on the data of the problem, Lemma 5.38,
- finite element error estimate, Theorem 5.41, which is based on the discrete inf-sup condition, and its discussion.

In summary, it will turn out that the numerical analysis can be performed for the velocity in a norm that is stronger than the norm used in the Galerkin method. However, for the stabilized method, it is difficult to incorporate the pressure appropriately in the analysis. A new idea in the finite element error analysis, compared with the Stokes equations, consists in using a discrete inf-sup condition for the complete bilinear form as a starting point. □

Remark 5.24 (SUPG/PSPG/grad-div (spg) Method) The SUPG/PSPG/grad-div method for solving the Oseen problem (5.2) is obtained by adding both a control of the strong residual of the momentum equation and the strong residual of the continuity equation on each mesh cell to the Galerkin finite element formulation (5.13). This classical residual-based stabilization reads as follows: Given $\boldsymbol{f} \in L^2(\Omega)$ and a triangulation $\mathcal{T}^h$ of Ω, find $\left(\boldsymbol{u}^h, p^h\right) \in V^h \times Q^h$ such that

$$A_{\text{spg}}\left(\left(\boldsymbol{u}^h, p^h\right), \left(\boldsymbol{v}^h, q^h\right)\right) = L_{\text{spg}}\left(\left(\boldsymbol{v}^h, q^h\right)\right) \quad \forall \left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h, \tag{5.33}$$

where the bilinear form $A_{\text{spg}} : \left(V \times \tilde{Q}\right) \times \left(V \times \tilde{Q}\right) \to \mathbb{R}$ is given by

$$\begin{aligned} & A_{\text{spg}}\left((\boldsymbol{u}, p), (\boldsymbol{v}, q)\right) \\ & = \nu\left(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}\right) + \left((\boldsymbol{b} \cdot \nabla)\boldsymbol{u} + c\boldsymbol{u}, \boldsymbol{v}\right) - \left(\nabla \cdot \boldsymbol{v}, p\right) + \left(\nabla \cdot \boldsymbol{u}, q\right) \\ & \quad + \sum_{K \in \mathcal{T}^h} \mu_K \left(\nabla \cdot \boldsymbol{u}, \nabla \cdot \boldsymbol{v}\right)_K + \sum_{E \in \mathcal{E}^h} \gamma_E \left([|p|]_E, [|q|]_E\right)_E \\ & \quad + \sum_{K \in \mathcal{T}^h} \left(-\nu\Delta\boldsymbol{u} + (\boldsymbol{b} \cdot \nabla)\boldsymbol{u} + c\boldsymbol{u} + \nabla p, \delta_K^v (\boldsymbol{b} \cdot \nabla)\boldsymbol{v} + \delta_K^p \nabla q\right)_K \end{aligned} \tag{5.34}$$

and the linear form $L_{\text{spg}} : \left(V \times \tilde{Q}\right) \to \mathbb{R}$ by

$$L_{\text{spg}}\left((\boldsymbol{v}, q)\right) = (\boldsymbol{f}, \boldsymbol{v}) + \sum_{K \in \mathcal{T}^h} \left(\boldsymbol{f}, \delta_K^v (\boldsymbol{b} \cdot \nabla)\boldsymbol{v} + \delta_K^p \nabla q\right)_K. \tag{5.35}$$

The correct asymptotic choice of the stabilization parameters $\mu_K \geq 0$ and γ_E, $\delta_K^v, \delta_K^p > 0$ in (5.34) and (5.35) will be determined by the finite element error analysis. The space $\tilde{Q}$ is defined in (4.122).

Individual equations for the velocity and the pressure test functions are obtained for the test functions $\left(\boldsymbol{v}^h, 0\right)$ and $\left(\boldsymbol{0}, -q^h\right)$: Find $\left(\boldsymbol{u}^h, p^h\right) \in V^h \times Q^h$ such that

$$\begin{aligned}
&\nu\left(\nabla \boldsymbol{u}^h, \nabla \boldsymbol{v}^h\right) + \left((\boldsymbol{b}\cdot\nabla)\,\boldsymbol{u}^h + c\boldsymbol{u}^h, \boldsymbol{v}^h\right) - \left(\nabla\cdot\boldsymbol{v}^h, p^h\right) \\
&\quad + \sum_{K\in\mathcal{T}^h} \mu_K \left(\nabla\cdot\boldsymbol{u}^h, \nabla\cdot\boldsymbol{v}^h\right)_K \\
&\quad + \sum_{K\in\mathcal{T}^h} \left(-\nu\Delta\boldsymbol{u}^h + (\boldsymbol{b}\cdot\nabla)\,\boldsymbol{u}^h + c\boldsymbol{u}^h + \nabla p^h, \delta_K^v (\boldsymbol{b}\cdot\nabla)\,\boldsymbol{v}^h\right)_K \\
&= \left(\boldsymbol{f}, \boldsymbol{v}^h\right) + \sum_{K\in\mathcal{T}^h} \left(\boldsymbol{f}, \delta_K^v (\boldsymbol{b}\cdot\nabla)\,\boldsymbol{v}^h\right)_K \quad \forall\ \boldsymbol{v}^h \in V^h,
\end{aligned}$$

and

$$\begin{aligned}
&-\left(\nabla\cdot\boldsymbol{u}^h, q^h\right) - \sum_{E\in\mathcal{E}^h} \gamma_E \left(\left[|p^h|\right]_E, \left[|q^h|\right]_E\right)_E \\
&\quad - \sum_{K\in\mathcal{T}^h} \left(-\nu\Delta\boldsymbol{u}^h + (\boldsymbol{b}\cdot\nabla)\,\boldsymbol{u}^h + c\boldsymbol{u}^h + \nabla p^h, \delta_K^p \nabla q^h\right)_K \\
&= -\sum_{K\in\mathcal{T}^h} \left(\boldsymbol{f}, \delta_K^p \nabla q^h\right)_K \quad \forall\ q^h \in Q^h.
\end{aligned} \tag{5.36}$$

The second equation shows that for $\delta_K^p > 0$ the finite element solution is generally not discretely divergence-free. But on the other hand, the grad-div term gives a control on the $L^2(\Omega)$ norm of the divergence.

Considering Dirichlet boundary conditions, as it is done in this section, an additive constant of the finite element pressure has to be fixed in (5.33). For this reason, $Q^h \subset L_0^2(\Omega)$ is still a correct condition for the finite element pressure space. □

Remark 5.25 (Comparison with the Prototype Stabilization (5.32)) In contrast to the prototype stabilization (5.32), the following modifications are contained in the SUPG/PSPG/grad-div (spg) method:

- The term $\nu\Delta\boldsymbol{v}^h$ is missing. A motivation for this modification is that the term is of minor importance if ν is small, which is the interesting case for stabilized methods.
- Likewise, the term $c\boldsymbol{v}^h$ does not appear. This term does not improve important properties of the discretization, like stability or the order of convergence.
- The term with the jumps of the finite element pressure appears. This term will be necessary for defining an appropriate norm in the case of discontinuous pressure approximations, see Lemma 5.33. If $Q^h \subset C(\Omega)$, then this term vanishes anyway.
- The velocity and pressure test functions in the stabilization term for the momentum balance may possess different weights. However, in practice these weights are chosen to be the same, see Remark 5.27.

□

Remark 5.26 (Role of the Different Stabilization Terms) The different stabilization terms have different functions:

- The stabilization term with the test function $(\boldsymbol{b} \cdot \nabla)\, \boldsymbol{v}$ is called Streamline-Upwind Petrov–Galerkin (SUPG) term. Sometimes, it is referred also as streamline diffusion (SD) term. This term stabilizes dominating convection.
- The pressure stabilization Petrov–Galerkin (PSPG) term is the stabilization term with the test function ∇q and for discontinuous discrete pressures the term with the pressure jumps across the faces. With these terms, a violation of the discrete inf-sup conditions (3.51) is stabilized, compare Sect. 4.5.1.
- Finally, the grad-div term is the stabilization term which involves the residual of the continuity equation. With this term, one gets an additional control on the violation of the conservation of mass, see Sect. 4.6.1.

□

Remark 5.27 (On the Stabilization Parameters) In the following, it will be assumed that $\delta_K = \delta_K^v = \delta_K^p$ for all $K \in \mathcal{T}^h$ and it is set

$$\delta = \max_{K \in \mathcal{T}^h} \delta_K, \quad \mu = \max_{K \in \mathcal{T}^h} \mu_K, \quad \gamma = \max_{E \in \mathcal{E}^h} \gamma_E. \tag{5.37}$$

□

Remark 5.28 (Stabilization Parameters with Dimensions) Considering the bilinear form (5.34) with dimensions, i.e., $\boldsymbol{u}, \boldsymbol{b}$ are velocities with $[\mathrm{m}/\mathrm{s}]$, p a pressure divided by density with $[\mathrm{m}^2/\mathrm{s}^2]$ and so on, one gets the following units for the stabilization parameters:

- $\mu_K : [\mathrm{m}^2/\mathrm{s}]$ – μ_K is a viscosity scale,
- $\gamma_E : [\mathrm{s}/\mathrm{m}]$ – γ_E^{-1} is a velocity scale,
- $\delta_K : [\mathrm{s}]$ – δ_K is a time scale.

□

Remark 5.29 (Historical Remarks to the SUPG Method) The SUPG method was introduced in Hughes and Brooks (1979), Brooks and Hughes (1982) for stabilizing convection-dominated convection-diffusion equations. Stabilizations of the Oseen equations and the stationary Navier–Stokes equations which contain the SUPG term were analyzed independently in Hansbo and Szepessy (1990), Lube and Tobiska (1990), Tobiska and Lube (1991), Franca and Frey (1992). A finite element error analysis of the SUPG/PSPG/grad-div method was given in Tobiska and Verfürth (1996) and another presentation can be found in Roos et al. (2008, Chap. IV.3.1).

□

Remark 5.30 (Inf-Sup Stable and Not Inf-Sup Stable Pairs of Finite Element Spaces) The SUPG/PSPG/grad-div method contains in particular the PSPG term. Hence, (5.33) is not a saddle point problem, compare Remark 4.92, and consequently, the inf-sup stability of the finite element spaces for velocity and pressure does not play a critical role if the SUPG/PSPG/grad-div method is used.

However, the finite element error analysis will reveal that the asymptotically optimal choice of the stabilization parameters is affected by the concrete choice of the pair of finite element spaces. For instance, the optimal choice of the stabilization parameters will be fundamentally different, e.g., for inf-sup stable pairs of finite element spaces and for equal order finite element spaces for velocity and pressure.

□

Lemma 5.31 (Consistency and Galerkin Orthogonality) *Let* $(\boldsymbol{u}, p) \in (V \cap H^2(\Omega)) \times (Q \cap H^1(\Omega))$ *be the solution of* (5.2) *and let* $\boldsymbol{b}, c \in L^\infty(\Omega), \boldsymbol{f} \in L^2(\Omega)$. *Consider conforming finite element spaces* V^h *and* Q^h, *then the SUPG/PSPG/grad-div method* (5.33) *is consistent, i.e.,*

$$A_{\mathrm{spg}}\left((\boldsymbol{u},p),\left(\boldsymbol{v}^h,q^h\right)\right) = L_{\mathrm{spg}}\left(\left(\boldsymbol{v}^h,q^h\right)\right), \quad \forall \left(\boldsymbol{v}^h,q^h\right) \in V^h \times Q^h, \tag{5.38}$$

and the Galerkin orthogonality (projection property)

$$A_{\mathrm{spg}}\left(\left(\boldsymbol{u}-\boldsymbol{u}^h,p-p^h\right),\left(\boldsymbol{v}^h,q^h\right)\right) = 0, \quad \forall \ \left(\boldsymbol{v}^h,q^h\right) \in V^h \times Q^h \tag{5.39}$$

holds.

Proof With the assumed regularity of the solution of (5.2), the local residuals in A_{spg} and L_{spg} belong to $L^2(K)$ for all $K \in \mathcal{T}^h$. The pressure does not possess jumps almost everywhere across the faces of the mesh cells. From the first property, it follows that the local residuals vanish for the solution of (5.2). Hence, all stabilization terms vanish if a sufficiently smooth solution of (5.2) is inserted in (5.33). The remaining terms are identical to Eq. (5.2).

The Galerkin orthogonality follows by subtracting (5.33) from (5.38). ■

Remark 5.32 (Norm for the First Part of the Analysis) Stabilized methods are analyzed generally in a norm which is connected to the stabilization, compare Sect. 4.5.1. This norm is in general stronger than the norm which is used in the analysis of the Galerkin finite element method. The norm used in the finite element error analysis for stabilized methods reveals which norms of the error are controlled by the stabilization.

For the SUPG/PSPG/grad-div method (5.33), the following norm will be used in the first part of the analysis

$$\begin{aligned}\|(\boldsymbol{v},q)\|_{\mathrm{spg}} = \Bigg(&\nu \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}^2 + \left\|c^{1/2}\boldsymbol{v}\right\|_{L^2(\Omega)}^2 + \sum_{K\in\mathcal{T}^h} \mu_K \|\nabla\cdot\boldsymbol{v}\|_{L^2(K)}^2 \\ &+ \sum_{E\in\mathcal{E}^h} \gamma_E \|[|q|]_E\|_{L^2(E)}^2 + \sum_{K\in\mathcal{T}^h} \delta_K \|(\boldsymbol{b}\cdot\nabla)\boldsymbol{v} + \nabla q\|_{L^2(K)}^2\Bigg)^{1/2}.\end{aligned} \tag{5.40}$$

This norm is mesh-dependent. For the Galerkin finite element discretization, only the first two terms of $\|\cdot\|_{\mathrm{spg}}$ are involved in the error bounds, see Theorem 5.14. The error estimation in the norm $\|(\cdot,\cdot)\|_{\mathrm{spg}}$ will give additional control on the error of the divergence, the term with the streamline derivative of the solution and the gradient of the pressure, and for discontinuous pressure approximations on the jumps of the discrete pressure. The amount of control depends on the parameters μ_K, δ_K, and γ_E, which have to be chosen such that

- the unique solvability of the finite element problem can be ensured and
- the asymptotic order of convergence of the error bounds becomes as high as possible.

□

Lemma 5.33 **($\|(\cdot,\cdot)\|_{\mathrm{spg}}$ Defines a Norm in $V^h \times Q^h$)** *Let $\delta_K > 0$ for all $K \in \mathcal{T}^h$, in the case $Q^h \not\subset H^1(\Omega)$ let $\gamma_E > 0$ for all $E \in \mathcal{E}^h$, and let V^h, Q^h be conforming finite element spaces. Then, $\|(\cdot,\cdot)\|_{\mathrm{spg}}$ defines a norm in $V^h \times Q^h$.*

Proof The proof can be performed in the same way as the proof of Lemma 4.94. ■

Theorem 5.34 (Existence and Uniqueness of a Solution of (5.33)) *Let $\boldsymbol{b}, c \in L^\infty(\Omega)$, let V^h and Q^h be conforming finite element spaces, and let the stabilization parameters be chosen such that*

$$0 < \delta_K \le \min\left\{\frac{h_K^2}{3\nu C_{\mathrm{inv}}^2}, \frac{1}{3\,\|c\|_{L^\infty(K)}}\right\}, \quad 0 \le \mu_K \le \mu < \infty, \tag{5.41}$$

and

$$0 < \gamma_E \le \gamma < \infty$$

if $Q^h \not\subset H^1(\Omega)$. Then, the finite element problem (5.33) possesses a unique solution.

Proof It will be shown that A_{spg} is coercive on $V^h \times Q^h$, that it is bounded on $V^h \times Q^h$, and that L_{spg} is bounded as well. Then, the statement of the theorem follows from the Theorem of Lax–Milgram, Theorem B.4.

First, coercivity is shown. It is

$$\begin{aligned} & A_{\mathrm{spg}}\left(\left(\boldsymbol{v}^h, q^h\right), \left(\boldsymbol{v}^h, q^h\right)\right) \\ &= \left\|\left(\boldsymbol{v}^h, q^h\right)\right\|_{\mathrm{spg}}^2 + \sum_{K\in\mathcal{T}^h} \delta_K \left(-\nu\Delta\boldsymbol{v}^h + c\boldsymbol{v}^h, (\boldsymbol{b}\cdot\nabla)\,\boldsymbol{v}^h + \nabla q^h\right)_K, \end{aligned} \tag{5.42}$$

where (5.4) was used. Now, the last terms on the right-hand side are estimated from above. One obtains with the Cauchy–Schwarz inequality (A.10), Young's inequality (A.5), and (5.41)

$$
\begin{aligned}
&\left|\sum_{K\in\mathcal{T}^h}\delta_K\left(c\boldsymbol{v}^h,(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{v}^h+\nabla q^h\right)_K\right| \\
&\le \frac{3}{2}\sum_{K\in\mathcal{T}^h}\delta_K\,\|c\|_{L^\infty(K)}\left\|c^{1/2}\boldsymbol{v}^h\right\|_{L^2(K)}^2+\frac{1}{6}\sum_{K\in\mathcal{T}^h}\delta_K\left\|(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{v}^h+\nabla q^h\right\|_{L^2(K)}^2 \\
&\le \frac{1}{2}\left\|c^{1/2}\boldsymbol{v}^h\right\|_{L^2(\Omega)}^2+\frac{1}{6}\sum_{K\in\mathcal{T}^h}\delta_K\left\|(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{v}^h+\nabla q^h\right\|_{L^2(K)}^2. \qquad (5.43)
\end{aligned}
$$

Using in addition the inverse inequality (C.35) yields

$$
\begin{aligned}
&\left|\sum_{K\in\mathcal{T}^h}\delta_K\left(-\nu\Delta\boldsymbol{v}^h,(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{v}^h+\nabla q^h\right)_K\right| \\
&\le \frac{3}{2}\sum_{K\in\mathcal{T}^h}\delta_K\frac{C_{\text{inv}}^2}{h_K^2}\nu^2\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(K)}^2+\frac{1}{6}\sum_{K\in\mathcal{T}^h}\delta_K\left\|(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{v}^h+\nabla q^h\right\|_{L^2(K)}^2 \\
&\le \frac{\nu}{2}\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)}^2+\frac{1}{6}\sum_{K\in\mathcal{T}^h}\delta_K\left\|(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{v}^h+\nabla q^h\right\|_{L^2(K)}^2. \qquad (5.44)
\end{aligned}
$$

Subtracting these upper bounds in (5.42) gives

$$
A_{\text{spg}}\left(\left(\boldsymbol{v}^h,q^h\right),\left(\boldsymbol{v}^h,q^h\right)\right)\ge\frac{1}{2}\left\|\left(\boldsymbol{v}^h,q^h\right)\right\|_{\text{spg}}^2,\quad\forall\ \left(\boldsymbol{v}^h,q^h\right)\in V^h\times Q^h, \qquad (5.45)
$$

which is the coercivity of A_{spg} on $V^h\times Q^h$ since $\left\|\left(\boldsymbol{v}^h,q^h\right)\right\|_{\text{spg}}$ defines a norm on $V^h\times Q^h$.

In the assumption of the Theorem of Lax–Milgram, the boundedness of the bilinear form and of the right-hand side is required. The bilinear form and the right-hand side are bounded if and only if they are continuous. Since all terms in A_{spg} and the right-hand side are defined with integrals, the continuity follows directly from the fact that all integrals in (5.34) and (5.35) are continuous. It is of course also possible to prove the continuity with direct estimates, e.g., using the tools that were employed for deriving estimate (4.128) for the PSPG method. ■

Lemma 5.35 (Stability of the Finite Element Solution) *Let the assumptions of Theorem 5.34 be satisfied, then the solution of* (5.33) *fulfills*

$$
\left\|\left(\boldsymbol{u}^h,p^h\right)\right\|_{\text{spg}}^2\le\frac{12}{5}\min\left\{\frac{\|\boldsymbol{f}\|_{H^{-1}(\Omega)}^2}{\nu},\frac{\|\boldsymbol{f}\|_{L_2(\Omega)}^2}{c_0}\right\}+4\sum_{K\in\mathcal{T}^h}\delta_K\,\|\boldsymbol{f}\|_{L^2(K)}^2. \qquad (5.46)
$$

If the finite element spaces fulfill the discrete inf-sup condition (3.51)*, then there holds the stability estimate*

$$\begin{aligned}
&\left\|p^h\right\|_{L^2(\Omega)} \qquad (5.47)\\
&\leq \frac{C}{\beta_{\text{is}}^h}\Bigg[\|\boldsymbol{f}\|_{H^{-1}(\Omega)} + \delta\,\|\boldsymbol{b}\|_{L^\infty(\Omega)}\,\|\boldsymbol{f}\|_{L^2(\Omega)} + \Bigg(\nu^{1/2} + \|c\|_{L^\infty(\Omega)}^{1/2} + \mu^{1/2}\\
&\quad + \|\boldsymbol{b}\|_{L^\infty(\Omega)}\min\left\{\frac{1}{\nu^{1/2}},\frac{1}{c_0^{1/2}}\right\} + \delta^{1/2}\,\|\boldsymbol{b}\|_{L^\infty(\Omega)}\Bigg)\left\|\left(\boldsymbol{u}^h,p^h\right)\right\|_{\text{spg}}\Bigg],
\end{aligned}$$

where the constant C does not depend on the data of the problem.

Proof As usual, the solution of the equations will be used as test function. Inserting $\left(\boldsymbol{u}^h,p^h\right)$ in (5.33) and using (5.4) gives

$$\begin{aligned}
\left\|\left(\boldsymbol{u}^h,p^h\right)\right\|_{\text{spg}}^2 &= \left(\boldsymbol{f},\boldsymbol{u}^h\right) + \sum_{K\in\mathcal{T}^h}\delta_K\left(\boldsymbol{f},(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{u}^h + \nabla p^h\right)_K\\
&\quad - \sum_{K\in\mathcal{T}^h}\delta_K\left(-\nu\Delta\boldsymbol{u}^h + c\boldsymbol{u}^h,(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{u}^h + \nabla p^h\right)_K.
\end{aligned}$$

The terms on the right-hand side are estimated by using the Cauchy–Schwarz inequality (A.10) and Young's inequality (A.5), see the proof of Lemma 5.8 for the first term and (5.43) and (5.44) for the other terms,

$$\begin{aligned}
\left|\left(\boldsymbol{f},\boldsymbol{u}^h\right)\right| &\leq \frac{3}{5\nu}\,\|\boldsymbol{f}\|_{H^{-1}(\Omega)}^2 + \frac{5\nu}{12}\left\|\nabla\boldsymbol{u}^h\right\|_{L^2(\Omega)}^2,\\
\left|\left(\boldsymbol{f},\boldsymbol{u}^h\right)\right| &\leq \frac{3}{5c_0}\,\|\boldsymbol{f}\|_{L^2(\Omega)}^2 + \frac{5}{12}\left\|c^{1/2}\boldsymbol{u}^h\right\|_{L^2(\Omega)}^2,
\end{aligned}$$

$$\begin{aligned}
&\left|\sum_{K\in\mathcal{T}^h}\delta_K\left(c\boldsymbol{u}^h,(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{u}^h + \nabla p^h\right)_K\right|\\
&\leq \frac{1}{3}\left\|c^{1/2}\boldsymbol{u}^h\right\|_{L^2(\Omega)}^2 + \frac{1}{4}\sum_{K\in\mathcal{T}^h}\delta_K\left\|(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{u}^h + \nabla p^h\right\|_{L^2(K)}^2,\\
&\left|\sum_{K\in\mathcal{T}^h}\delta_K\left(-\nu\Delta\boldsymbol{u}^h,(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{u}^h + \nabla p^h\right)_K\right|\\
&\leq \frac{\nu}{3}\left\|\nabla\boldsymbol{u}^h\right\|_{L^2(\Omega)}^2 + \frac{1}{4}\sum_{K\in\mathcal{T}^h}\delta_K\left\|(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{u}^h + \nabla p^h\right\|_{L^2(K)}^2.
\end{aligned}$$

For the remaining term, one obtains

$$\begin{aligned}&\left|\sum_{K\in\mathcal{T}^h}\delta_K\left(\boldsymbol{f},(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{u}^h+\nabla p^h\right)_K\right|\\&\le\sum_{K\in\mathcal{T}^h}\delta_K\,\|\boldsymbol{f}\|_{L^2(K)}^2+\frac{1}{4}\sum_{K\in\mathcal{T}^h}\delta_K\left\|(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{u}^h+\nabla p^h\right\|_{L^2(K)}^2.\end{aligned}$$

Absorbing all terms in the left-hand side gives the first statement of the lemma.

For inf-sup stable discretizations, the stability estimate for the finite element pressure starts with the discrete inf-sup condition (3.51), e.g., see the proof of Lemma 5.8. Then, $-(\nabla\cdot\boldsymbol{v}^h,p^h)$ is substituted by (5.33). In the next step, it is noted that the terms with q^h cancel, which follows from choosing in (5.33) as test functions $(\mathbf{0},q^h)$. For the terms coming from the stabilization, the Cauchy–Schwarz inequality for sums (A.2) is applied and the bounds for the stabilization parameters (5.41) are used to obtain, e.g.,

$$\begin{aligned}&\sum_{K\in\mathcal{T}^h}\delta_K\frac{C_{\mathrm{inv}}}{h_K}\nu\left\|\nabla\boldsymbol{u}^h\right\|_{L^2(K)}\|\boldsymbol{b}\|_{L^\infty(K)}\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(K)}\\&\le C\delta^{1/2}\,\|\boldsymbol{b}\|_{L^\infty(\Omega)}\,\nu^{1/2}\left\|\nabla\boldsymbol{u}^h\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)}.\end{aligned}$$

In this way, and by using the Poincaré–Friedrichs inequality (A.12), one gets a factor $\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)}$ in all terms of the numerator such that these terms cancel with the denominator. ■

Remark 5.36 (On the Stability Estimate)

- The estimate (5.46) is in a stronger norm than for the Galerkin discretization, see (5.14).
- An estimate of $\left\|p^h\right\|_{L^2(\Omega)}$ for inf-sup stable pairs of finite element spaces in terms of the data of the problem is obtained by inserting (5.46) in the right-hand side of (5.47).

□

Remark 5.37 (On the Way of Proving an Error Estimate in a Norm Which Contains the $L^2(\Omega)$ Norm of the Pressure) The approach that was used in Tobiska and Verfürth (1996) to prove an error estimate for the SUPG/PSPG/grad-div method differs from the approach for the Galerkin discretization. This way utilizes a discrete inf-sup condition for the bilinear form A_{spg}.

So far, the norm $\|\cdot\|_{\mathrm{spg}}$ was used in the analysis. This norm does not control the $L^2(\Omega)$ norm of the pressure. If $Q^h\subset H^1(\Omega)$, only the gradient of the pressure is contained in a mixed term with the streamline derivative of the velocity. However, from the analysis of the continuous problem, it is known that the natural norm for the pressure is the $L^2(\Omega)$ norm. The finite element error analysis which will be performed now uses a norm which, in addition to $\|\cdot\|_{\text{spg}}$, possesses a contribution of the pressure

$$\|(\boldsymbol{v}, q)\|_{\text{spg},p} = \left(\|(\boldsymbol{v}, q)\|_{\text{spg}}^2 + \omega_{\text{pres}}^{-2} \|q\|_{L^2(\Omega)}^2\right)^{1/2} \tag{5.48}$$

with

$$\omega_{\text{pres}} = \max\left\{1, \nu^{-1/2}, \|c\|_{L^\infty(\Omega)}^{1/2}\right\}. \tag{5.49}$$

However, for the interesting cases of small ν and large c (small time steps), the contribution of the pressure becomes small:

$$\omega_{\text{pres}}^{-2} = \min\left\{1, \nu, \|c\|_{L^\infty(\Omega)}^{-1}\right\}.$$

□

Lemma 5.38 (Inf-Sup Condition for A_{spg}) *Let the assumptions of Theorem 5.34 be satisfied, let in addition exist a positive constants δ_0 such that for all triangulations in the family $\{\mathcal{T}^h\}_{h>0}$ it holds*

$$0 < \delta_0 h_K^2 \le \delta_K \quad \forall\, K \in \mathcal{T}^h. \tag{5.50}$$

If $Q^h \not\subset H^1(\Omega)$, then let there be a positive constant γ_0 such that for all triangulations in the family $\{\mathcal{T}^h\}_{h>0}$ one has

$$0 < \gamma_0 h_E \le \gamma_E \quad \forall\, E \in \mathcal{E}^h. \tag{5.51}$$

Then, there is a constant $\beta_{\text{spg}}^h > 0$, such that

$$\inf_{\substack{(\boldsymbol{v}^h, q^h) \in V^h \times Q^h \\ \|(\boldsymbol{v}^h, q^h)\|_{\text{spg},p} = 1}} \sup_{\substack{(\boldsymbol{w}^h, r^h) \in V^h \times Q^h \\ \|(\boldsymbol{w}^h, r^h)\|_{\text{spg},p} = 1}} A_{\text{spg}}\left(\left(\boldsymbol{v}^h, q^h\right), \left(\boldsymbol{w}^h, r^h\right)\right) \ge \beta_{\text{spg}}^h. \tag{5.52}$$

The inf-sup constant is independent of h and ν, see also Remark 5.39.

Proof The plan of the proof is as follows:

i) For all $\left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h$ a function $\boldsymbol{w}^h \in V^h$ will be constructed such that one obtains an estimate of the form

$$A_{\text{spg}}\left(\left(\boldsymbol{v}^h, q^h\right), \left(\boldsymbol{w}^h, 0\right)\right) \ge \text{negative terms} + C \left\|q^h\right\|_{L^2(\Omega)}^2$$

with $C > 0$.

ii) A linear combination of the result of the first step and the coercivity condition (5.45) is constructed such that the negative terms from the first step are absorbed. The result is of the form

$$A_{\text{spg}}\left(\left(\boldsymbol{v}^h, q^h\right), \left((1-\rho)\boldsymbol{v}^h + \rho\boldsymbol{w}^h, (1-\rho)q^h\right)\right) \geq C \left\|\left(\boldsymbol{v}^h, q^h\right)\right\|_{\text{spg},p}^2$$

with $C > 0$ and $\rho \in (0, 1)$.

iii) It is proved that

$$\left\|\left((1-\rho)\boldsymbol{v}^h + \rho\boldsymbol{w}^h, (1-\rho)q^h\right)\right\|_{\text{spg},p} \leq 2 \left\|\boldsymbol{v}^h, q^h\right\|_{\text{spg},p},$$

which is inserted in the right-hand side of the result from the second step. Then, the discrete inf-sup condition (5.52) follows with straightforward arguments.

Step i). Consider an arbitrary pair $\left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h$ and set for abbreviation

$$X := \left(\sum_{K\in\mathcal{T}^h} \delta_K \left\|(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{v}^h + \nabla q^h\right\|_{L^2(K)}^2\right)^{1/2},$$

$$Y := \left(\sum_{K\in\mathcal{T}^h} \mu_K \left\|\nabla\cdot\boldsymbol{v}^h\right\|_{L^2(K)}^2\right)^{1/2},$$

$$Z := \left(\sum_{E\in\mathcal{E}^h} \gamma_E \left\|\left[\left|q^h\right|\right]_E\right\|_{L^2(E)}^2\right)^{1/2}.$$

From (5.45), it is known that

$$A_{\text{spg}}\left(\left(\boldsymbol{v}^h, q^h\right), \left(\boldsymbol{v}^h, q^h\right)\right) \geq \frac{1}{2}\left\|\left(\boldsymbol{v}^h, q^h\right)\right\|_{\text{spg}}^2, \quad \forall\ \left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h. \tag{5.53}$$

The spaces V and Q satisfy the inf-sup condition, see Theorem 3.46. From Corollary 3.44, it follows that for $q^h \subset Q$ there is a $\boldsymbol{w} \in V_{\text{div}}^{\perp} \subset V$ with $\nabla\cdot\boldsymbol{w} = -q^h$ and $\|\nabla\boldsymbol{w}\|_{L^2(\Omega)} \leq C\left\|q^h\right\|_{L^2(\Omega)}$. Let $\boldsymbol{w}^h = I^h\boldsymbol{w}$ be an interpolation of $\boldsymbol{w}$ that fulfills the standard interpolation properties. In particular, the following interpolation properties are assumed

$$\left\|\boldsymbol{w} - I^h\boldsymbol{w}\right\|_{H^m(K)} \leq C h_K^{l-m}\|\boldsymbol{w}\|_{H^l(K)} \quad \forall\ \boldsymbol{w} \in H^l(K), \tag{5.54}$$

$$\left\|\boldsymbol{w} - I^h\boldsymbol{w}\right\|_{L^2(E)} \leq C h_E^{l-1/2}\|\boldsymbol{w}\|_{H^l(K_1\cup K_2)} \quad \forall\ \boldsymbol{w} \in H^l(K_1\cup K_2), \tag{5.55}$$

where E is the joint face of K_1 and K_2, $H^0(K) = L^2(K)$, $0 \leq m \leq 2$, $\max\{1, m\} \leq l$. In Roos et al. (2008, Chap. IV.3.1), the interpolation operator proposed in Scott and Zhang (1990) is used. It follows with $l = m = 1$ in (5.54) that

$$\begin{aligned}\left\|\nabla\boldsymbol{w}^h\right\|_{L^2(\Omega)} &\leq \left\|\nabla\boldsymbol{w} - \nabla\boldsymbol{w}^h\right\|_{L^2(\Omega)} + \|\nabla\boldsymbol{w}\|_{L^2(\Omega)} \leq C\|\nabla\boldsymbol{w}\|_{L^2(\Omega)} + \|\nabla\boldsymbol{w}\|_{L^2(\Omega)}\\ &\leq C_1\left\|q^h\right\|_{L^2(\Omega)}.\end{aligned} \tag{5.56}$$

Adding $-\left(\nabla \cdot \boldsymbol{w}, q^h\right) + \left(\nabla \cdot \boldsymbol{w}, q^h\right)$ and using integration by parts yields

$$
\begin{aligned}
& A_{\text{spg}}\left(\left(\boldsymbol{v}^h, q^h\right), \left(\boldsymbol{w}^h, 0\right)\right) \\
&= \nu\left(\nabla \boldsymbol{v}^h, \nabla \boldsymbol{w}^h\right) + \left((\boldsymbol{b}\cdot\nabla)\,\boldsymbol{v}^h, \boldsymbol{w}^h\right) + \left(c\boldsymbol{v}^h, \boldsymbol{w}^h\right) - \left(\nabla\cdot\boldsymbol{w}, q^h\right) \\
&\quad - \sum_{K\in\mathcal{T}^h}\left(\boldsymbol{w}-\boldsymbol{w}^h, \nabla q^h\right)_K + \sum_{E\in\mathcal{E}^h}\left(\left(\boldsymbol{w}-\boldsymbol{w}^h\right)\cdot\boldsymbol{n}_E, \left[|q^h|\right]_E\right)_E \\
&\quad + \sum_{K\in\mathcal{T}^h}\mu_K\left(\nabla\cdot\boldsymbol{v}^h, \nabla\cdot\boldsymbol{w}^h\right)_K \\
&\quad + \sum_{K\in\mathcal{T}^h}\delta_K\left(-\nu\Delta\boldsymbol{v}^h + (\boldsymbol{b}\cdot\nabla)\,\boldsymbol{v}^h + c\boldsymbol{v}^h + \nabla q^h, (\boldsymbol{b}\cdot\nabla)\,\boldsymbol{w}^h\right)_K.
\end{aligned}
$$

Now, applying the Cauchy–Schwarz inequality (A.10), adding $(\boldsymbol{b}\cdot\nabla)\boldsymbol{v}^h - (\boldsymbol{b}\cdot\nabla)\boldsymbol{v}^h$, using the triangle inequality, the Poincaré–Friedrichs inequality (A.12), (3.41), (5.56), and $\nabla\cdot\boldsymbol{w} = -q^h$ leads to

$$
\begin{aligned}
& A_{\text{spg}}\left(\left(\boldsymbol{v}^h, q^h\right), \left(\boldsymbol{w}^h, 0\right)\right) \\
&\ge -C_1\Big(\nu\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)} + \|\boldsymbol{b}\|_{L^\infty(\Omega)}\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)} + \|c\|_{L^\infty(\Omega)}^{1/2}\left\|c^{1/2}\boldsymbol{v}^h\right\|_{L^2(\Omega)}\Big)\left\|q^h\right\|_{L^2(\Omega)} \\
&\quad - \sum_{K\in\mathcal{T}^h}\left\|\boldsymbol{w}-\boldsymbol{w}^h\right\|_{L^2(K)}\Big(\left\|(\boldsymbol{b}\cdot\nabla)\boldsymbol{v}^h + \nabla q^h\right\|_{L^2(K)} + \left\|(\boldsymbol{b}\cdot\nabla)\boldsymbol{v}^h\right\|_{L^2(K)}\Big) \\
&\quad - \sum_{E\in\mathcal{E}^h}\left\|\boldsymbol{w}-\boldsymbol{w}^h\right\|_{L^2(E)}\left\|\left[|q^h|\right]_E\right\|_{L^2(E)} \\
&\quad - \sum_{K\in\mathcal{T}^h}\mu_K\left\|\nabla\cdot\boldsymbol{v}^h\right\|_{L^2(K)}\left\|\nabla\boldsymbol{w}^h\right\|_{L^2(K)} \\
&\quad - \sum_{K\in\mathcal{T}^h}\delta_K\Bigg[\nu\left\|\Delta\boldsymbol{v}^h\right\|_{L^2(K)} + \left\|c^{1/2}\right\|_{L^\infty(K)}\left\|c^{1/2}\boldsymbol{v}^h\right\|_{L^2(K)} \\
&\quad + \left\|(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{v}^h + \nabla q^h\right\|_{L^2(K)}\Bigg]\|\boldsymbol{b}\|_{L^\infty(K)}\left\|\nabla\boldsymbol{w}^h\right\|_{L^2(K)} + \left\|q^h\right\|_{L^2(\Omega)}^2.
\end{aligned}
$$

The individual terms are estimated separately, always using the Cauchy–Schwarz inequality, with the goal to get the factor $\left\|q^h\right\|_{L^2(\Omega)}$. One obtains with (5.54), (5.50), the Poincaré–Friedrichs inequality, and (5.56)

$$
\begin{aligned}
& \sum_{K\in\mathcal{T}^h}\left\|\boldsymbol{w}-\boldsymbol{w}^h\right\|_{L^2(K)}\Big(\left\|(\boldsymbol{b}\cdot\nabla)\boldsymbol{v}^h + \nabla q^h\right\|_{L^2(K)} + \left\|(\boldsymbol{b}\cdot\nabla)\boldsymbol{v}^h\right\|_{L^2(K)}\Big) \\
&\quad \le C\sum_{K\in\mathcal{T}^h}h_K\|\boldsymbol{w}\|_{H^1(K)}\Big(\left\|(\boldsymbol{b}\cdot\nabla)\boldsymbol{v}^h + \nabla q^h\right\|_{L^2(K)} + \left\|(\boldsymbol{b}\cdot\nabla)\boldsymbol{v}^h\right\|_{L^2(K)}\Big)
\end{aligned}
$$

$$
\begin{aligned}
&\leq CX \left(\sum_{K \in \mathcal{T}^h} h_K^2 \delta_K^{-1} \|\boldsymbol{w}\|_{H^1(K)}^2 \right)^{1/2} \\
&\quad + C \|\boldsymbol{b}\|_{L^\infty(\Omega)} \left(\sum_{K \in \mathcal{T}^h} h_K^2 \left\|\nabla \boldsymbol{v}^h\right\|_{L^2(K)}^2 \right)^{1/2} \left(\sum_{K \in \mathcal{T}^h} \|\boldsymbol{w}\|_{H^1(K)}^2 \right)^{1/2} \\
&\leq C \left(\delta_0^{-1/2} X + h \|\boldsymbol{b}\|_{L^\infty(\Omega)} \left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)} \right) \|\nabla \boldsymbol{w}\|_{L^2(\Omega)} \\
&\leq CC_1 \left(\delta_0^{-1/2} X + h \|\boldsymbol{b}\|_{L^\infty(\Omega)} \left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)} \right) \left\|q^h\right\|_{L^2(\Omega)} .
\end{aligned}
$$

For the next term, one gets with (5.54), (5.51), the Poincaré–Friedrichs inequality, and (5.56)

$$
\begin{aligned}
&\sum_{E \in \mathcal{E}^h} \left\|\boldsymbol{w} - \boldsymbol{w}^h\right\|_{L^2(E)} \left\|[|q^h|]_E\right\|_{L^2(E)} \\
&\leq C \sum_{E \in \mathcal{E}^h} \|\boldsymbol{w}\|_{H^1(K_1 \cup K_2)} \left\|[|q^h|]_E\right\|_{L^2(E)} \leq CZ \left(\sum_{E \in \mathcal{E}^h} \gamma_E^{-1} h_E \|\boldsymbol{w}\|_{H^1(K_1 \cup K_2)}^2 \right)^{1/2} \\
&\leq CC_1 \gamma_0^{-1/2} Z \left\|q^h\right\|_{L^2(\Omega)} .
\end{aligned}
$$

The estimate of the next term uses again (5.56), leading to

$$
\sum_{K \in \mathcal{T}^h} \mu_K \left\|\nabla \cdot \boldsymbol{v}^h\right\|_{L^2(K)} \left\|\nabla \boldsymbol{w}^h\right\|_{L^2(K)} \leq C_1 \mu^{1/2} Y \left\|q^h\right\|_{L^2(\Omega)} .
$$

For the SUPG/PSPG term, one obtains with the inverse estimate (C.35), (5.41), and (5.56)

$$
\begin{aligned}
&\sum_{K \in \mathcal{T}^h} \delta_K \nu \left\|\Delta \boldsymbol{v}^h\right\|_{L^2(K)} \|\boldsymbol{b}\|_{L^\infty(K)} \left\|\nabla \boldsymbol{w}^h\right\|_{L^2(K)} \\
&\leq \sum_{K \in \mathcal{T}^h} C_{\mathrm{inv}} \delta_K h_K^{-1} \nu \left\|\nabla \boldsymbol{v}^h\right\|_{L^2(K)} \|\boldsymbol{b}\|_{L^\infty(K)} \left\|\nabla \boldsymbol{w}^h\right\|_{L^2(K)} \\
&\leq CC_1 h \|\boldsymbol{b}\|_{L^\infty(\Omega)} \left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)} \left\|q^h\right\|_{L^2(\Omega)} .
\end{aligned}
$$

Similarly, one gets

$$
\begin{aligned}
&\sum_{K \in \mathcal{T}^h} \delta_K \|c\|_{L^\infty(K)}^{1/2} \left\|c^{1/2} \boldsymbol{v}^h\right\|_{L^2(K)} \|\boldsymbol{b}\|_{L^\infty(K)} \left\|\nabla \boldsymbol{w}^h\right\|_{L^2(K)} \\
&\leq C_1 \delta \|c\|_{L^\infty(\Omega)}^{1/2} \|\boldsymbol{b}\|_{L^\infty(\Omega)} \left\|c^{1/2} \boldsymbol{v}^h\right\|_{L^2(\Omega)} \left\|q^h\right\|_{L^2(\Omega)}
\end{aligned}
$$

and

$$\sum_{K\in\mathcal{T}^h} \delta_K \left\| (\boldsymbol{b}\cdot\nabla)\,\boldsymbol{v}^h + \nabla q^h \right\|_{L^2(K)} \|\boldsymbol{b}\|_{L^\infty(K)} \left\|\nabla \boldsymbol{w}^h\right\|_{L^2(K)}$$
$$\leq C_1\delta^{1/2} \|\boldsymbol{b}\|_{L^\infty(\Omega)} X \left\|q^h\right\|_{L^2(\Omega)}.$$

Collecting terms and applying Young's inequality (A.5) term by term leads to

$$\begin{aligned}
&A_{\mathrm{spg}}\left(\left(\boldsymbol{v}^h, q^h\right), \left(\boldsymbol{w}^h, 0\right)\right)\\
&\geq -\Bigg[\left(C_1\frac{\nu}{\nu^{1/2}} + C_1\frac{\|\boldsymbol{b}\|_{L^\infty(\Omega)}}{\nu^{1/2}} + CC_1h\frac{\|\boldsymbol{b}\|_{L^\infty(\Omega)}}{\nu^{1/2}}\right)\nu^{1/2}\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)}\\
&\quad+\left(C_1\|c\|_{L^\infty(\Omega)}^{1/2} + C_1\delta\|c\|_{L^\infty(\Omega)}^{1/2}\|\boldsymbol{b}\|_{L^\infty(\Omega)}\right)\left\|c^{1/2}\boldsymbol{v}^h\right\|_{L^2(\Omega)}\\
&\quad+C_1\mu^{1/2}Y + CC_1\gamma_0^{-1/2}Z + \left(CC_1\delta_0^{-1/2} + C_1\delta^{1/2}\|\boldsymbol{b}\|_{L^\infty(\Omega)}\right)X\Bigg]\left\|q^h\right\|_{L^2(\Omega)}\\
&\quad+\left\|q^h\right\|_{L^2(\Omega)}^2\\
&\geq -\omega_{\mathrm{pres}}\widehat{C}\left[\nu^{1/2}\|\nabla\boldsymbol{v}\|_{L^2(\Omega)} + \left\|c^{1/2}\boldsymbol{v}^h\right\|_{L^2(\Omega)} + Y + Z + X\right]\left\|q^h\right\|_{L^2(\Omega)} + \left\|q^h\right\|_{L^2(\Omega)}^2\\
&\geq -2\omega_{\mathrm{pres}}^2\widehat{C}^2\left\|\left(\boldsymbol{v}^h, q^h\right)\right\|_{\mathrm{spg}}^2 + \frac{3}{8}\left\|q^h\right\|_{L^2(\Omega)}^2,
\end{aligned}\tag{5.57}$$

where ω_{pres} is defined in (5.49) and

$$\begin{aligned}
\widehat{C} = \max\Big\{&C_1\nu + C_1\|\boldsymbol{b}\|_{L^\infty(\Omega)} + CC_1h\|\boldsymbol{b}\|_{L^\infty(\Omega)},\\
&C_1 + C_1\delta\|\boldsymbol{b}\|_{L^\infty(\Omega)}, C_1\mu^{1/2}, CC_1\gamma_0^{-1/2}, \left(CC_1\delta_0^{-1/2} + C_1\delta^{1/2}\|\boldsymbol{b}\|_{L^\infty(\Omega)}\right)\Big\}.
\end{aligned}\tag{5.58}$$

Step ii). Let $\rho > 0$, then adding $(1-\rho)$ times (5.53) and ρ times (5.57) gives

$$\begin{aligned}
&A_{\mathrm{spg}}\left(\left(\boldsymbol{v}^h, q^h\right), \left((1-\rho)\boldsymbol{v}^h + \rho\boldsymbol{w}^h, (1-\rho)q^h\right)\right)\\
&\geq \left(\frac{1}{2} - \frac{\rho}{2} - 2\rho\left(\omega_{\mathrm{pres}}\widehat{C}\right)^2\right)\left\|\left(\boldsymbol{v}^h, q^h\right)\right\|_{\mathrm{spg}}^2 + \frac{3\rho}{8}\left\|q^h\right\|_{L^2(\Omega)}^2.
\end{aligned}$$

The term in the parentheses should be positive. Requiring that

$$\frac{1}{2} - \frac{\rho}{2} - 2\rho\left(\omega_{\mathrm{pres}}\widehat{C}\right)^2 \leq \frac{3\rho}{8}\omega_{\mathrm{pres}}^2 \tag{5.59}$$

leads to

$$\frac{1}{2} \le \left(\frac{3}{8}\omega_{\mathrm{pres}}^2 + \frac{1}{2} + 2\left(\omega_{\mathrm{pres}}\widehat{C}\right)^2\right)\rho \le \left(\frac{3}{8} + \frac{1}{2} + 2\widehat{C}^2\right)\omega_{\mathrm{pres}}^2\rho$$

because $\omega_{\mathrm{pres}} \ge 1$. The following choice satisfies the requirement (5.59)

$$0 < \rho = \frac{4}{7 + 16\widehat{C}^2}\omega_{\mathrm{pres}}^{-2} < 1. \tag{5.60}$$

For the term in parentheses, one gets with this choice and $\omega_{\mathrm{pres}}^{-2} \le 1$

$$\begin{aligned}\frac{1}{2} - \frac{\rho}{2} - 2\rho\left(\omega_{\mathrm{pres}}\widehat{C}\right)^2 &= \frac{1}{2} - \frac{2}{7 + 16\widehat{C}^2}\left(\omega_{\mathrm{pres}}^{-2} + 4\widehat{C}^2\right) \\ &\ge \frac{1}{2} - \frac{2 + 8\widehat{C}^2}{7 + 16\widehat{C}^2} = \frac{1}{2}\left(1 - \frac{4 + 16\widehat{C}^2}{7 + 16\widehat{C}^2}\right) > 0.\end{aligned}$$

In this way, one obtains the estimate

$$\begin{aligned}&A_{\mathrm{spg}}\left(\left(\boldsymbol{v}^h, q^h\right), \left((1-\rho)\boldsymbol{v}^h + \rho\boldsymbol{w}^h, (1-\rho)q^h\right)\right) \\ &\ge \frac{3}{14 + 32\widehat{C}^2}\left(\left\|\left(\boldsymbol{v}^h, q^h\right)\right\|_{\mathrm{spg}}^2 + \omega_{\mathrm{pres}}^{-2}\left\|q^h\right\|_{L^2(\Omega)}^2\right) \\ &= \frac{3}{14 + 32\widehat{C}^2}\left\|\left(\boldsymbol{v}^h, q^h\right)\right\|_{\mathrm{spg},p}^2.\end{aligned} \tag{5.61}$$

Step iii). One has by the triangle inequality, the definition of $\|(\cdot,\cdot)\|_{\mathrm{spg},p}$, (3.41), (5.56), (5.49), (5.58), (5.60), and once more the definition of $\|(\cdot,\cdot)\|_{\mathrm{spg},p}$

$$\begin{aligned}&\left\|\left((1-\rho)\boldsymbol{v}^h + \rho\boldsymbol{w}^h, (1-\rho)q^h\right)\right\|_{\mathrm{spg},p} \\ &\le (1-\rho)\left\|\boldsymbol{v}^h, q^h\right\|_{\mathrm{spg},p} + \rho\Bigg(\nu\left\|\nabla\boldsymbol{w}^h\right\|_{L^2(\Omega)}^2 + \nu\left\|c^{1/2}\boldsymbol{w}^h\right\|_{L^2(\Omega)}^2 \\ &\quad + \sum_{K\in\mathcal{T}^h}\mu_K\left\|\nabla\cdot\boldsymbol{w}^h\right\|_{L^2(K)}^2 + \sum_{K\in\mathcal{T}^h}\delta_K\left\|(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{w}^h\right\|_{L^2(K)}^2\Bigg)^{1/2} \\ &\le \left\|\boldsymbol{v}^h, q^h\right\|_{\mathrm{spg},p} + \rho\left(C_1^2\nu + C_1^2\|c\|_{L^\infty(\Omega)} + C_1^2\mu d + C_1^2\delta\|\boldsymbol{b}\|_{L^\infty(\Omega)}^2\right)^{1/2}\left\|q^h\right\|_{L^2(\Omega)} \\ &\le \left\|\boldsymbol{v}^h, q^h\right\|_{\mathrm{spg},p} + 2\rho\max\left\{1, \nu^{-1/2}, \|c\|_{L^\infty(\Omega)}^{1/2}\right\} \\ &\quad \times\max\left\{C_1\nu, C_1, C_1\mu^{1/2}, C_1\delta^{1/2}\|\boldsymbol{b}\|_{L^\infty(\Omega)}\right\}\left\|q^h\right\|_{L^2(\Omega)}\end{aligned}$$

$$
\begin{aligned}
&\le \left\| \boldsymbol{v}^h, q^h \right\|_{\mathrm{spg},p} + 2\rho\omega_{\mathrm{pres}}\widehat{C} \left\| q^h \right\|_{L^2(\Omega)} \\
&\le \left\| \boldsymbol{v}^h, q^h \right\|_{\mathrm{spg},p} + \frac{8\widehat{C}}{7+16\widehat{C}^2}\omega_{\mathrm{pres}}^{-1} \left\| q^h \right\|_{L^2(\Omega)} \\
&\le (1+1) \left\| \boldsymbol{v}^h, q^h \right\|_{\mathrm{spg},p} = 2 \left\| \boldsymbol{v}^h, q^h \right\|_{\mathrm{spg},p},
\end{aligned}
$$

where $8\widehat{C}/(7 + 16\widehat{C}^2) \le 1/\sqrt{7} < 1$ has been used in the last step.

Inserting this estimate in (5.61) gives for all $\left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h$

$$
\begin{aligned}
&A_{\mathrm{spg}} \left(\left(\boldsymbol{v}^h, q^h\right), \left((1-\rho)\boldsymbol{v}^h + \rho\boldsymbol{w}^h, (1-\rho)q^h\right) \right) \\
&\ge \frac{3}{28 + 64\widehat{C}^2} \left\| \left(\boldsymbol{v}^h, q^h\right) \right\|_{\mathrm{spg},p} \left\| \left((1-\rho)\boldsymbol{v}^h + \rho\boldsymbol{w}^h, (1-\rho)q^h\right) \right\|_{\mathrm{spg},p}
\end{aligned}
$$

or

$$
\begin{aligned}
A_{\mathrm{spg}} &\left(\frac{\left(\boldsymbol{v}^h, q^h\right)}{\|(\boldsymbol{v}^h, q^h)\|_{\mathrm{spg},p}}, \frac{\left((1-\rho)\boldsymbol{v}^h + \rho\boldsymbol{w}^h, (1-\rho)q^h\right)}{\|((1-\rho)\boldsymbol{v}^h + \rho\boldsymbol{w}^h, (1-\rho)q^h)\|_{\mathrm{spg},p}} \right) \\
&\ge \frac{3}{28 + 64\widehat{C}^2} = \beta_{\mathrm{spg}}^h. \qquad (5.62)
\end{aligned}
$$

The arguments of the bilinear form are normalized with respect to $\|\cdot\|_{\mathrm{spg},p}$. The inequality stays valid if, for each $\left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h$, the supremum of normalized functions with respect to the second argument of the bilinear form is considered

$$
\sup_{\substack{(\boldsymbol{w}^h, r^h) \in V^h \times Q^h \\ \|(\boldsymbol{w}^h, r^h)\|_{\mathrm{spg},p} = 1}} A_{\mathrm{spg}} \left(\frac{\left(\boldsymbol{v}^h, q^h\right)}{\|(\boldsymbol{v}^h, q^h)\|_{\mathrm{spg},p}}, \left(\boldsymbol{w}^h, r^h\right) \right) \ge \beta_{\mathrm{spg}}^h.
$$

Since this inequality holds still for all $\left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h$, it is valid also for the infimum of normalized functions with respect to the first argument, such that (5.52) is proved. ■

Remark 5.39 (On β_{spg}^h)

- From the assumptions on the scaling of the Oseen equations, see Remark 5.3, it follows that the constant $\widehat{C}$ in (5.58) behaves like $\mathcal{O}(1)$ with respect to the coefficients of the Oseen problem. Therefore, one obtains from (5.62) also that $\beta_{\mathrm{spg}}^h = \mathcal{O}(1)$ with respect to the coefficients of the Oseen equations.
- An inf-sup condition of form (5.52) does not hold for the Galerkin discretization. That means, one has to expect that $\beta_{\mathrm{spg}}^h \to 0$ if the stabilization terms vanish. The vanishing of the SUPG/PSPG term is described by the size of δ_0 from (5.50). It

can be seen in (5.58) that $\widehat{C} = \mathcal{O}\left(\delta_0^{-1/2}\right)$ from what follows that $\beta_{\mathrm{spg}}^h = \mathcal{O}(\delta_0)$ for $\delta_0 \to 0$.

- If $Z \neq 0$, then one obtains with the same arguments that $\beta_{\mathrm{spg}}^h = \mathcal{O}(\gamma_0)$ for $\gamma_0 \to 0$.

□

Corollary 5.40 (Existence and Uniqueness of a Solution of (5.33)) *Let the assumptions of Lemma 5.38 be satisfied, then problem (5.33) possesses a unique solution.*

Proof The statement of the corollary is a direct consequence of the inf-sup condition (5.52) and Lemma B.15. ■

Theorem 5.41 (Error Estimate) *Let the assumptions of Lemma 5.38 be satisfied, let $(\boldsymbol{u}, p)$ be the solution of (5.2), let a quasi-uniform family of triangulations of the domain be given with the finite element spaces*

$$P_k \text{ or } Q_k \subseteq V^h \subset V, \quad k \geq 1, \qquad P_l \text{ or } Q_l \subseteq Q^h \subset Q, \quad l \geq 0,$$

and let $\left(\boldsymbol{u}^h, p^h\right) \in V^h \times Q^h$ be the solution of (5.33). With the assumptions $\boldsymbol{u} \in H^{k+1}(\Omega)$ and $p \in H^{l+1}(\Omega)$, the following error estimate holds

$$\begin{aligned}
&\left\| \left(\boldsymbol{u} - \boldsymbol{u}^h, p - p^h\right) \right\|_{\mathrm{spg},p} \\
&\leq C\Bigg[h^k \Bigg(\nu^{1/2} + \left(h + \delta^{1/2} h\right) \|c\|_{L^\infty(\Omega)}^{1/2} + \delta^{1/2} \|\boldsymbol{b}\|_{L^\infty(\Omega)}^{1/2} + \delta^{1/2} \\
&\quad + \delta_0^{-1/2} + \gamma_0^{-1/2} + \mu^{1/2} \Bigg) \|\boldsymbol{u}\|_{H^{k+1}(\Omega)} \\
&\quad + h^l \Bigg(\delta^{1/2} + h \min\left\{ \nu^{-1/2}, \max_{K \in \mathcal{T}^h} \left\{ \mu_K^{-1/2} \right\} \right\} + h\omega_{\mathrm{pres}}^{-1} \\
&\quad + \gamma^{1/2} \left(h + h^{1/2}\right) \Bigg) \|p\|_{H^{l+1}(\Omega)} \Bigg] \qquad (5.63)
\end{aligned}$$

with C independent of the coefficients of the problem. The terms $\gamma_0^{-1/2}$ and $\gamma^{1/2}\left(h + h^{1/2}\right)$ are not present if $Q^h \subset H^1(\Omega)$.

Proof Let $\boldsymbol{v}^h = I^h \boldsymbol{u}$ be the Lagrangian interpolation of $\boldsymbol{u}$ in V^h and $q^h = P_{L^2}^h q$ be the $L^2(\Omega)$ projection of q onto $P_l \subseteq Q^h$ (or $Q_l \subseteq Q^h$).

The triangle inequality gives

$$\left\| \left(\boldsymbol{u} - \boldsymbol{u}^h, p - p^h\right) \right\|_{\mathrm{spg},p} \leq \left\| \left(\boldsymbol{u} - \boldsymbol{v}^h, p - q^h\right) \right\|_{\mathrm{spg},p} + \left\| \left(\boldsymbol{u}^h - \boldsymbol{v}^h, p^h - q^h\right) \right\|_{\mathrm{spg},p}. \qquad (5.64)$$

First, the interpolation error will be considered. From the choice of q^h, it follows that the term with the pressure jumps vanishes for $l \geq 1$. Using the definition of the norm, the interpolation estimates (C.14), estimate (3.41), the estimates for the $L^2(\Omega)$ projection (C.28) and (C.29), estimate (C.34) for $l = 0$, and (5.37) yields

$$
\begin{aligned}
&\left\| \left(\boldsymbol{u} - \boldsymbol{v}^h, p - q^h\right) \right\|_{\mathrm{spg},p} \\
&= \Bigg(\nu \left\| \nabla \left(\boldsymbol{u} - \boldsymbol{v}^h\right) \right\|_{L^2(\Omega)}^2 + \left\| c^{1/2} \left(\boldsymbol{u} - \boldsymbol{v}^h\right) \right\|_{L^2(\Omega)}^2 + \sum_{K \in \mathcal{T}^h} \mu_K \left\| \nabla \cdot \left(\boldsymbol{u} - \boldsymbol{v}^h\right) \right\|_{L^2(K)}^2 \\
&\quad + \sum_{E \in \mathcal{E}^h} \gamma_E \left\| \left[\left| p - q^h \right|\right]_E \right\|_{L^2(E)}^2 + \sum_{K \in \mathcal{T}^h} \delta_K \left\| (\boldsymbol{b} \cdot \nabla) \left(\boldsymbol{u} - \boldsymbol{v}^h\right) + \nabla \left(p - q^h\right) \right\|_{L^2(K)}^2 \\
&\quad + \omega_{\mathrm{pres}}^{-2} \left\| p - q^h \right\|_{L^2(\Omega)}^2 \Bigg)^{1/2} \\
&\leq C \Bigg(\nu h^{2k} \|\boldsymbol{u}\|_{H^{k+1}(\Omega)}^2 + \|c\|_{L^\infty(\Omega)} h^{2k+2} \|\boldsymbol{u}\|_{H^{k+1}(\Omega)}^2 + \mu h^{2k} \|\boldsymbol{u}\|_{H^{k+1}(\Omega)}^2 \\
&\quad + \gamma h^{2l+1} \|p\|_{H^{l+1}(\Omega)}^2 + \delta \|\boldsymbol{b}\|_{L^\infty(\Omega)} h^{2k} \|\boldsymbol{u}\|_{H^{k+1}(\Omega)}^2 + \delta h^{2l} \|p\|_{H^{l+1}(\Omega)}^2 \\
&\quad + \omega_{\mathrm{pres}}^{-2} h^{2l+2} \|p\|_{H^{l+1}(\Omega)}^2 \Bigg)^{1/2} \\
&\leq C \left(\nu + h^2 \|c\|_{L^\infty(\Omega)} + \mu + \delta \|\boldsymbol{b}\|_{L^\infty(\Omega)}\right)^{1/2} h^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)} \\
&\quad + C \left(\gamma h + \delta + \omega_{\mathrm{pres}}^{-2} h^2\right)^{1/2} h^l \|p\|_{H^{l+1}(\Omega)} \\
&\leq C \left(\nu^{1/2} + h \|c\|_{L^\infty(\Omega)}^{1/2} + \mu^{1/2} + \delta^{1/2} \|\boldsymbol{b}\|_{L^\infty(\Omega)}^{1/2}\right) h^k \|\boldsymbol{u}\|_{H^{k+1}(\Omega)} \\
&\quad + C \left(\gamma^{1/2} h^{1/2} + \delta^{1/2} + \omega_{\mathrm{pres}}^{-1} h\right) h^l \|p\|_{H^{l+1}(\Omega)} .
\end{aligned}
\tag{5.65}
$$

Next, the second term of (5.64) will be estimated. Scaling the inf-sup condition (5.52) and using the Galerkin orthogonality (5.39) leads to

$$
\begin{aligned}
&\left\| \left(\boldsymbol{u}^h - \boldsymbol{v}^h, p^h - q^h\right) \right\|_{\mathrm{spg},p} \\
&\leq \frac{1}{\beta_{\mathrm{spg}}^h} \sup_{\substack{\left(\boldsymbol{w}^h, r^h\right) \in V^h \times Q^h \\ \left\| \left(\boldsymbol{w}^h, r^h\right) \right\|_{\mathrm{spg},p} = 1}} A_{\mathrm{spg}} \left(\left(\boldsymbol{u}^h - \boldsymbol{v}^h, p^h - q^h\right), \left(\boldsymbol{w}^h, r^h\right) \right) \\
&= \frac{1}{\beta_{\mathrm{spg}}^h} \sup_{\substack{\left(\boldsymbol{w}^h, r^h\right) \in V^h \times Q^h \\ \left\| \left(\boldsymbol{w}^h, r^h\right) \right\|_{\mathrm{spg},p} = 1}} A_{\mathrm{spg}} \left(\left(\boldsymbol{u} - \boldsymbol{v}^h, p - q^h\right), \left(\boldsymbol{w}^h, r^h\right) \right) .
\end{aligned}
\tag{5.66}
$$

With this step, one got rid of $\left(\boldsymbol{u}^h, p^h\right)$ in the estimate. Now, all terms on the right-hand side of (5.66) are estimate individually. The goal is to bound the terms with $\left(\boldsymbol{w}^h, r^h\right)$ with terms that are included in $\left\|\left(\boldsymbol{w}^h, r^h\right)\right\|_{\text{spg},p}$ and then to estimate this norm with 1.

With the Cauchy–Schwarz inequality (A.10), $\left\|\left(\boldsymbol{w}^h, r^h\right)\right\|_{\text{spg},p} = 1$, and the interpolation estimate (C.14), one obtains

$$\begin{aligned}\nu\left(\nabla\left(\boldsymbol{u}-\boldsymbol{v}^h\right), \nabla \boldsymbol{w}^h\right) &\le \nu\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{v}^h\right)\right\|_{L^2(\Omega)}\left\|\nabla \boldsymbol{w}^h\right\|_{L^2(\Omega)} \\ &\le C\nu^{1/2}h^k\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}\left\|\left(\boldsymbol{w}^h, r^h\right)\right\|_{\text{spg},p} \\ &= C\nu^{1/2}h^k\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} \qquad (5.67)\end{aligned}$$

and

$$\begin{aligned}\left(c\left(\boldsymbol{u}-\boldsymbol{v}^h\right), \boldsymbol{w}^h\right) &\le \left\|c^{1/2}\left(\boldsymbol{u}-\boldsymbol{v}^h\right)\right\|_{L^2(\Omega)}\left\|c^{1/2}\boldsymbol{w}^h\right\|_{L^2(\Omega)} \\ &\le C\|c\|_{L^\infty(\Omega)}^{1/2}h^{k+1}\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}. \qquad (5.68)\end{aligned}$$

The estimate of the next term starts with integration by parts. Then, $\nabla\cdot\boldsymbol{b} = 0$ is used, the Cauchy–Schwarz inequalities for integrals (A.10) and for sums (A.2) are applied, $\left\|\left(\boldsymbol{w}^h, r^h\right)\right\|_{\text{spg},p} = 1$ is utilized, the conditions (5.50) and (5.51) on the stabilization parameters are used, the interpolation estimate (C.14), the estimate (C.15), and $\left\|\left(\boldsymbol{w}^h, r^h\right)\right\|_{\text{spg},p} = 1$ are applied to get

$$\begin{aligned}&\left(\boldsymbol{b}\cdot\nabla\left(\boldsymbol{u}-\boldsymbol{v}^h\right), \boldsymbol{w}^h\right) + \left(\nabla\cdot\left(\boldsymbol{u}-\boldsymbol{v}^h\right), r^h\right) \\ &= -\left(\boldsymbol{u}-\boldsymbol{v}^h, (\boldsymbol{b}\cdot\nabla)\boldsymbol{w}^h\right) - \left(\boldsymbol{u}-\boldsymbol{v}^h, \nabla r^h\right) + \sum_{K\in\mathcal{T}^h}\sum_{E\subset\partial K}\left(\left(\boldsymbol{u}-\boldsymbol{v}^h\right)\cdot\boldsymbol{n}_E, r^h\right)_E \\ &= -\left(\boldsymbol{u}-\boldsymbol{v}^h, (\boldsymbol{b}\cdot\nabla)\boldsymbol{w}^h + \nabla r^h\right) + \sum_{E\in\mathcal{E}^h}\left(\left(\boldsymbol{u}-\boldsymbol{v}^h\right)\cdot\boldsymbol{n}_E, [|r^h|]_E\right)_E \\ &\le \left(\sum_{K\in\mathcal{T}^h}\delta_K^{-1}\left\|\boldsymbol{u}-\boldsymbol{v}^h\right\|_{L^2(K)}^2\right)^{1/2}\left(\sum_{K\in\mathcal{T}^h}\delta_K\left\|(\boldsymbol{b}\cdot\nabla)\boldsymbol{w}^h + \nabla r^h\right\|_{L^2(K)}^2\right)^{1/2} \\ &\quad + \left(\sum_{E\in\mathcal{E}^h}\gamma_E^{-1}\left\|\boldsymbol{u}-\boldsymbol{v}^h\right\|_{L^2(E)}^2\right)^{1/2}\left(\sum_{E\in\mathcal{E}^h}\gamma_E\left\|[|r^h|]_E\right\|_{L^2(E)}^2\right)^{1/2} \\ &\le \left(C\delta_0^{-1}\sum_{K\in\mathcal{T}^h}h_K^{-2}h_K^{2k+2}\|\boldsymbol{u}\|_{H^{k+1}(K)}^2\right)^{1/2}\end{aligned}$$

$$
+\left(C\gamma_0^{-1}\sum_{E\in\mathcal{E}^h} h_E^{-1}h_E^{2k+1}\left\|\boldsymbol{u}\right\|^2_{H^{k+1}(K_1\cup K_2)}\right)^{1/2}
$$

$$
\leq C\left(\delta_0^{-1/2}+\gamma_0^{-1/2}\right)h^k\left\|\boldsymbol{u}\right\|_{H^{k+1}(\Omega)}, \tag{5.69}
$$

where $E = K_1 \cap K_2$. The next part of the right-hand side of (5.66) can be estimated with the Cauchy–Schwarz inequality, (3.41), the norm property of $(\boldsymbol{w}^h, r^h)$, and the projection estimate (C.28)

$$
\begin{aligned}
\left(\nabla\cdot\boldsymbol{w}^h, p-q^h\right) &\leq \left\|p-q^h\right\|_{L^2(\Omega)}\left\|\nabla\cdot\boldsymbol{w}^h\right\|_{L^2(\Omega)}\\
&\leq \left\|p-q^h\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{w}^h\right\|_{L^2(\Omega)}\\
&= \nu^{-1/2}\left\|p-q^h\right\|_{L^2(\Omega)}\nu^{1/2}\left\|\nabla\boldsymbol{w}^h\right\|_{L^2(\Omega)}\\
&\leq \nu^{-1/2}\left\|p-q^h\right\|_{L^2(\Omega)}\\
&\leq C\nu^{-1/2}h^{l+1}\left\|p\right\|_{H^{l+1}(\Omega)}.
\end{aligned}
$$

For bounding this term, also a different part of $\left\|\left(\boldsymbol{w}^h, r^h\right)\right\|_{\mathrm{spg},p}$ can be utilized, which gives

$$
\begin{aligned}
\left(\nabla\cdot\boldsymbol{w}^h, p-q^h\right) &\leq \left\|p-q^h\right\|_{L^2(\Omega)}\left\|\nabla\cdot\boldsymbol{w}^h\right\|_{L^2(\Omega)}\\
&\leq \max_{K\in\mathcal{T}^h}\left\{\mu_K^{-1/2}\right\}\left\|p-q^h\right\|_{L^2(\Omega)}\sum_{K\in\mathcal{T}^h}\mu_K^{1/2}\left\|\nabla\cdot\boldsymbol{w}^h\right\|_{L^2(K)}\\
&\leq \max_{K\in\mathcal{T}^h}\left\{\mu_K^{-1/2}\right\}\left\|p-q^h\right\|_{L^2(\Omega)}\left(\sum_{K\in\mathcal{T}^h}\mu_K\left\|\nabla\cdot\boldsymbol{w}^h\right\|^2_{L^2(K)}\right)^{1/2}\\
&\leq \max_{K\in\mathcal{T}^h}\left\{\mu_K^{-1/2}\right\}\left\|p-q^h\right\|_{L^2(\Omega)}\\
&\leq C\max_{K\in\mathcal{T}^h}\left\{\mu_K^{-1/2}\right\}h^{l+1}\left\|p\right\|_{H^{l+1}(\Omega)}.
\end{aligned}
$$

Altogether, one obtains

$$
\left(\nabla\cdot\boldsymbol{w}^h, p-q^h\right) \leq C\min\left\{\nu^{-1/2}, \max_{K\in\mathcal{T}^h}\left\{\mu_K^{-1/2}\right\}\right\}h^{l+1}\left\|p\right\|_{H^{l+1}(\Omega)}. \tag{5.70}
$$

The term with the pressure jumps vanishes for $l \geq 1$ since $\left[\left|p-q^h\right|\right]_E = 0$ for all faces because of the choice of q^h. For $l = 0$, one gets with (C.34)

$$
\sum_{E\in\mathcal{E}^h}\gamma_E\left(\left[\left|p-q^h\right|\right]_E, \left[\left|r^h\right|\right]_E\right)_E \leq C\gamma^{1/2}h^{l+1}\left\|p\right\|_{H^{l+1}(\Omega)}. \tag{5.71}
$$

Finally, the SUPG/PSPG terms have to be considered. One obtains with the Cauchy–Schwarz inequality for integrals and for sums, $\left\|\left(\boldsymbol{w}^h, r^h\right)\right\|_{\mathrm{spg},p} = 1$, the condition (5.41) on the stabilization parameter, and the interpolation estimate (C.14)

$$
\begin{aligned}
&\sum_{K\in\mathcal{T}^h} \delta_K \left(-\nu\Delta\left(\boldsymbol{u}-\boldsymbol{v}^h\right) + c\left(\boldsymbol{u}-\boldsymbol{v}^h\right), \left(\boldsymbol{b}\cdot\nabla\right)\boldsymbol{w}^h + \nabla r^h\right)_K \\
&\le \sum_{K\in\mathcal{T}^h} \left(\nu\delta_K^{1/2}\left\|\Delta\left(\boldsymbol{u}-\boldsymbol{v}^h\right)\right\|_{L^2(K)} + \|c\|_{L^\infty(K)}\delta_K^{1/2}\left\|\boldsymbol{u}-\boldsymbol{v}^h\right\|_{L^2(K)}\right) \\
&\quad \times\delta_K^{1/2}\left\|\left(\boldsymbol{b}\cdot\nabla\right)\boldsymbol{w}^h + \nabla r^h\right\|_{L^2(K)} \\
&\le \left[\left(\sum_{K\in\mathcal{T}^h} \nu^2\delta_K\left\|\Delta\left(\boldsymbol{u}-\boldsymbol{v}^h\right)\right\|_{L^2(K)}^2\right)^{1/2}\right. \\
&\quad \left.+\left(\sum_{K\in\mathcal{T}^h}\|c\|_{L^\infty(K)}^2\delta_K\left\|\boldsymbol{u}-\boldsymbol{v}^h\right\|_{L^2(K)}^2\right)^{1/2}\right] \\
&\quad \times\left(\sum_{K\in\mathcal{T}^h}\delta_K\left\|\left(\boldsymbol{b}\cdot\nabla\right)\boldsymbol{w}^h + \nabla r^h\right\|_{L^2(K)}^2\right)^{1/2} \\
&\le C\left(\sum_{K\in\mathcal{T}^h}\nu h_K^2\left\|\Delta\left(\boldsymbol{u}-\boldsymbol{v}^h\right)\right\|_{L^2(K)}^2\right)^{1/2} + \delta^{1/2}\|c\|_{L^\infty(\Omega)}\left\|\boldsymbol{u}-\boldsymbol{v}^h\right\|_{L^2(\Omega)} \\
&\le C\nu^{1/2}hh^{k-1}\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + C\delta^{1/2}\|c\|_{L^\infty(\Omega)}h^{k+1}\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} \\
&\le C\left(\nu^{1/2} + \delta^{1/2}\|c\|_{L^\infty(\Omega)}h\right)h^k\|\boldsymbol{u}\|_{H^{k+1}(\Omega)}.
\end{aligned}
\tag{5.72}
$$

In a similar way, one gets

$$
\begin{aligned}
&\sum_{K\in\mathcal{T}^h} \delta_K\left(\left(\boldsymbol{b}\cdot\nabla\right)\left(\boldsymbol{u}-\boldsymbol{v}^h\right) + \nabla\left(p-q^h\right), \left(\boldsymbol{b}\cdot\nabla\right)\boldsymbol{w}^h + \nabla r^h\right)_K \\
&\le \left[\left(\sum_{K\in\mathcal{T}^h}\delta_K\|\boldsymbol{b}\|_{L^\infty(K)}^2\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{v}^h\right)\right\|_{L^2(K)}^2\right)^{1/2}\right. \\
&\quad \left.+\left(\sum_{K\in\mathcal{T}^h}\delta_K\left\|\nabla\left(p-q^h\right)\right\|_{L^2(K)}^2\right)^{1/2}\right]\left(\sum_{K\in\mathcal{T}^h}\delta_K\left\|\left(\boldsymbol{b}\cdot\nabla\right)\boldsymbol{w}^h + \nabla r^h\right\|_{L^2(K)}^2\right)^{1/2} \\
&\le C\delta^{1/2}\|\boldsymbol{b}\|_{L^\infty(\Omega)}h^k\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + C\delta^{1/2}h^l\|p\|_{H^{l+1}(\Omega)}.
\end{aligned}
\tag{5.73}
$$

Collecting the estimates (5.65) and (5.67)–(5.73) gives the error bound (5.63). ■

Remark 5.42 (Optimal Asymptotics for the Stabilization Parameters) Based on the error estimate (5.63), one tries to determine the asymptotic form of the stabilization parameters such that the error bound becomes asymptotically optimal. However, one should be aware that changes of the stabilization parameters also change the norm on the left-hand side of (5.63). Thus, small stabilization parameters provide only a weak control on certain individual norms of $\|(\boldsymbol{u}-\boldsymbol{u}^h, p-p^h)\|spg,p$.

Let $\{\mathcal{T}^h\}$ be a family of uniform triangulations.

- Consider first the case $\nu < h$. In this case, it is more precise to speak of order of error reduction than of order of convergence. It will be assumed that $\mu \ge \nu$ such that

$$\min\left\{\nu^{-1/2}, \mu^{-1/2}\right\} = \mu^{-1/2}.$$

  - *Inf-sup stable pairs with* $k = l+1$. The optimal order of error reduction which can be achieved is k. Hence, the term in the first parentheses of the right-hand side of (5.63) should be $\mathcal{O}(1)$ and the term in the second parentheses $\mathcal{O}(h)$. Concentrating on the most important terms, one has to calibrate

$$\delta^{1/2}, \mu^{1/2}, \frac{\delta^{1/2}}{h}, \frac{1}{\mu^{1/2}}, \frac{\gamma^{1/2}}{h^{1/2}},$$

    where the last three terms come from the second parentheses, which has to be scaled by h^{-1} such that the order of error reduction becomes k. It can be seen that one has to choose $\delta \sim h^2$, $\mu \sim 1$, and, if $Q^h \not\subset H^1(\Omega)$ then $\gamma \sim h$.
  - *Equal-order pairs with* $k = l \ge 1$. In this case, the terms in both parentheses on the right-hand side of (5.63) should scale the same way. The most important terms apart of $\delta_0^{-1/2}$ and $\gamma_0^{-1/2}$ are

$$\delta^{1/2}, \mu^{1/2}, \frac{h}{\mu^{1/2}}, \gamma^{1/2}h^{1/2}.$$

    Thus, one gets $\mu \sim h$ such that $\mu^{1/2} \sim h^{1/2}$ and $h/\mu^{1/2} \sim h^{1/2}$. Choosing in addition $\delta \sim h$ and $\gamma \sim 1$, all terms in the parentheses, apart of $\delta_0^{-1/2}$ and $\gamma_0^{-1/2}$, are of order $k + 1/2$. In view of (5.50), one can think of $\delta_0 \sim h^{-1}$ if $\delta \sim h$ such that $\delta_0^{-1/2} \sim h^{1/2}$. Similarly, one has from (5.51) that $\gamma_0 \sim h^{-1}$ if $\gamma \sim 1$ such that $\gamma_0^{-1/2} \sim h^{1/2}$. Altogether, one can expect an error reduction of order $k + 1/2$.

  The obtained parameter choices do not contradict the assumption $\mu \ge \nu$.
- Consider now the case $\nu \ge h$. The first parentheses can be at best constant with respect to h because of the term $\nu^{1/2}$, such that the optimal order of convergence is k.

- *Inf-sup stable pairs with* $k = l + 1$. The situation is similar as in the case $\nu < h$, only for μ there is more freedom. Since choosing μ large leads to a strong norm on the left-hand side of (5.63), it makes sense to choose the same asymptotic scalings of the stabilization parameters like for $\nu < h$.
- *Equal-order pairs with* $k = l \geq 1$. It is sufficient that both parentheses are constant with respect to h. There is now some freedom for choosing the parameters. Small parameters lead to a lower bound on the right-hand side of (5.63) and large parameters to stronger norms on the left-hand side of (5.63). To get a small error bound, within the assumptions of the numerical analysis, like (5.50) and (5.51), one has to choose $\delta \sim h^2$ and $\gamma \sim h$. Note that these asymptotic choices correspond to the parameter choices (4.130) of the PSPG method for the Stokes equations.

  For the grad-div parameter $\mu \sim h$ is still a correct choice, in particular if ν is small, but for large ν also other choices are possible, like $\mu \sim 1$. □

Example 5.43 (Analytical Example Which Supports the Error Estimate (5.63)) The same problem as in Example 5.18 is considered. Results for the inf-sup stable Q_2/Q_1 Taylor–Hood finite element and for the equal-order pair P_1/P_1 will be presented. For both pairs it is $Q^h \subset H^1(\Omega)$ such that $\gamma_E = 0$ and the terms with γ_0 and γ do not appear on the right-hand side of (5.63). Note that the norm $\|\cdot\|_{\mathrm{spg}}$ changes if the coefficients ν and c of the problem are changed.

The Q_2/Q_1 Finite Element For the Q_2/Q_1 finite element, the optimal asymptotic choice of stabilization parameters is $\delta_K = \mathcal{O}\left(h_K^2\right)$ and $\mu_K = \mathcal{O}(1)$. Numerical studies in Matthies et al. (2009) investigated the sensitivity of the errors for wide ranges of parameters for several examples. From these studies, it can be concluded that $\mu_K = 0.2$ is a good choice. The SUPG/PSPG parameter is not that important. It only should be not too large. The results presented in Figs. 5.6, 5.7, and 5.8 were obtained with $\delta_K = 0.1 h_K^2$.

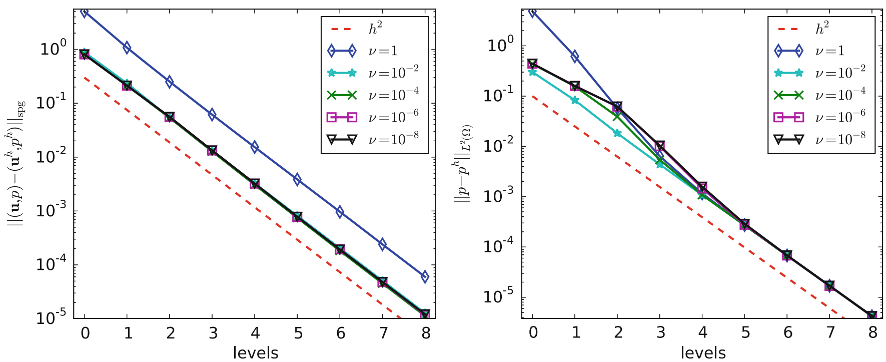

Fig. 5.6 Example 5.43. SUPG/PSPG/grad-div method with Q_2/Q_1, convergence of the errors $\left\|(\boldsymbol{u}, p) - (\boldsymbol{u}^h, p^h)\right\|_{\mathrm{spg}}$ and $\left\|p - p^h\right\|_{L^2(\Omega)}$ for $c = 0$ and different values of ν

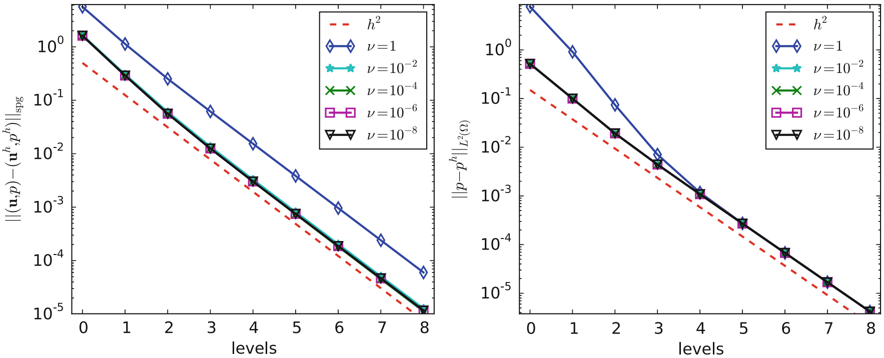

Fig. 5.7 Example 5.43. SUPG/PSPG/grad-div method with Q_2/Q_1, convergence of the errors $\left\|(\boldsymbol{u}, p) - (\boldsymbol{u}^h, p^h)\right\|_{\text{spg}}$ and $\left\|p - p^h\right\|_{L^2(\Omega)}$ for $c = 100$ and different values of ν

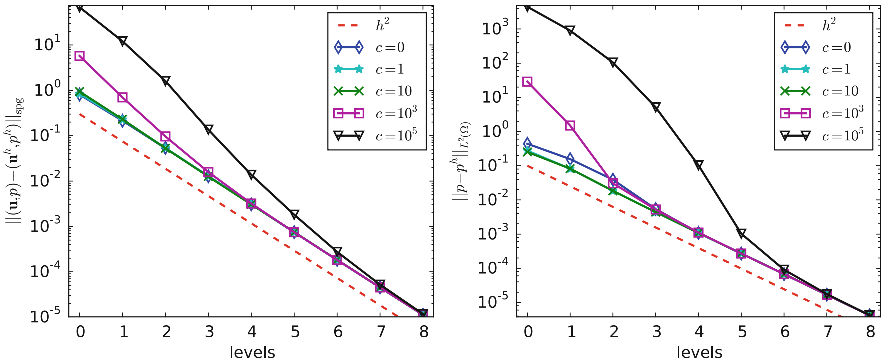

Fig. 5.8 Example 5.43. SUPG/PSPG/grad-div method with Q_2/Q_1, convergence of the errors $\left\|(\boldsymbol{u}, p) - (\boldsymbol{u}^h, p^h)\right\|_{\text{spg}}$ and $\left\|p - p^h\right\|_{L^2(\Omega)}$ for $\nu = 10^{-4}$ and different values of c

The norm on the left-hand side of the error estimate (5.63) can be split into two parts, namely into $\left\|\left(\boldsymbol{u} - \boldsymbol{u}^h, p - p^h\right)\right\|_{\text{spg}}$ and $\left\|p - p^h\right\|_{L^2(\Omega)}$. Both parts will be studied separately. Second order reduction of the error in the norm $\|\cdot\|_{\text{spg}}$ was predicted by the numerical analysis. This order can be clearly seen in Figs. 5.6, 5.7, and 5.8 for different combinations of the coefficients of the Oseen equations. The errors do not depend on the inverse of ν. In addition, one can see also a second order error reduction for the pressure error in $L^2(\Omega)$. The bound for this error is the right-hand side of (5.63) scaled with $\nu^{-1/2}$. However, a dependency of the error on ν can be observed only on coarse grids. In Fig. 5.8, larger errors on coarse grids for large values of c can be noticed.

P_1/P_1 Finite Element Remark 5.42 shows that the optimal stabilization parameters for the P_1/P_1 finite element are $\delta_K = \mu_K = \mathcal{O}(h_K)$ on coarse grids, i.e., if $\nu < h_K$, and $\delta_K = \mathcal{O}\left(h_K^2\right)$ else. For the grad-div parameter, the analysis does not lead to a concrete asymptotic behavior. To prevent a sharp change of this parameter, the

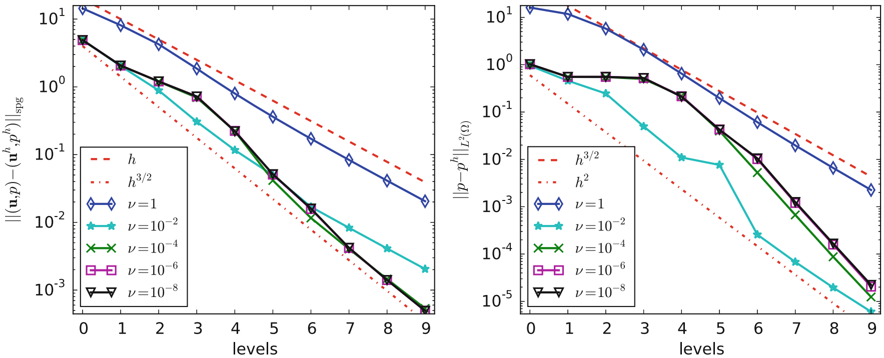

Fig. 5.9 Example 5.43. SUPG/PSPG/grad-div method with P_1/P_1, convergence of the errors $\left\|(\boldsymbol{u},p)-(\boldsymbol{u}^h,p^h)\right\|_{\text{spg}}$ and $\left\|p-p^h\right\|_{L^2(\Omega)}$ for $c=0$ and different values of ν

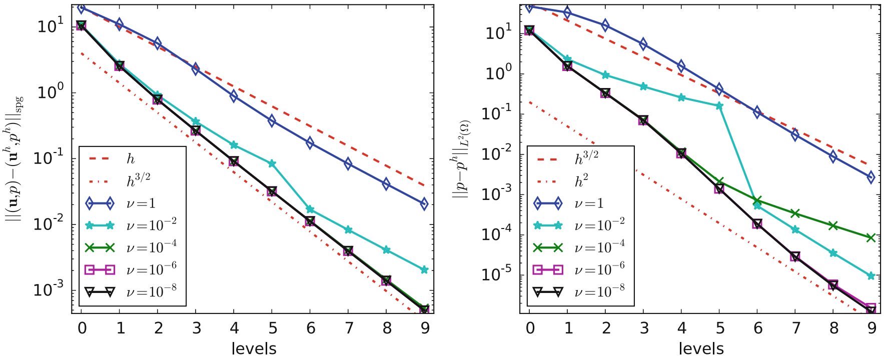

Fig. 5.10 Example 5.43. SUPG/PSPG/grad-div method with P_1/P_1, convergence of the errors $\left\|(\boldsymbol{u},p)-(\boldsymbol{u}^h,p^h)\right\|_{\text{spg}}$ and $\left\|p-p^h\right\|_{L^2(\Omega)}$ for $c=100$ and different values of ν

results presented in Figs. 5.9, 5.10, and 5.11 were computed with the same grad-div parameter for all situations. The concrete choices of the stabilization parameters were

$$\delta_K=\begin{cases}0.5h_K & \text{if } \nu<h_K,\\ 0.5h_K^2 & \text{else,}\end{cases}\qquad \mu_K=0.5h_K. \tag{5.74}$$

The predictions of the numerical analysis are an error reduction in $\|\cdot\|_{\text{spg}}$ of order 1.5 on coarse grids, i.e., if $\nu<h_K$, and a convergence of the error in $\|\cdot\|_{\text{spg}}$ of order 1 on fine grids, where fine has to be understood with respect to ν. Both predictions are supported by the results presented in Figs. 5.9, 5.10, and 5.11. The errors do not depend on the inverse of the viscosity.

Additionally, the error of the pressure in $L^2(\Omega)$ is presented in Figs. 5.9, 5.10, and 5.11. Often, the order of error reduction is better than predicted from the theory.

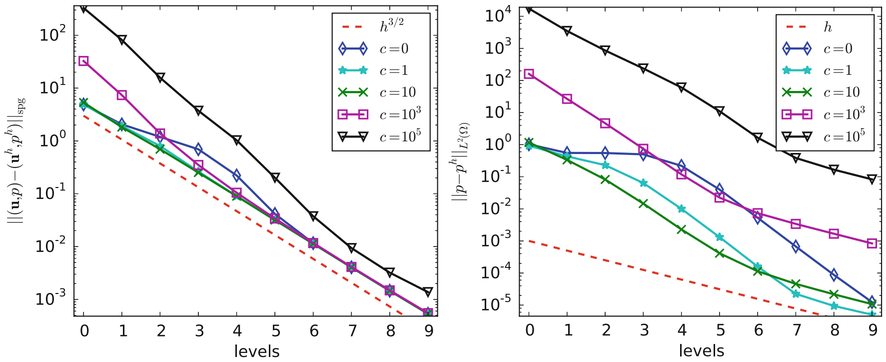

Fig. 5.11 Example 5.43. SUPG/PSPG/grad-div method with P_1/P_1, convergence of the errors $\left\|(\boldsymbol{u},p)-(\boldsymbol{u}^h,p^h)\right\|_{\mathrm{spg}}$ and $\left\|p-p^h\right\|_{L^2(\Omega)}$ for $\nu = 10^{-4}$ and different values of c

Small values of the viscosity give in this example smaller pressure errors. The kink in the curves for $\nu = 10^{-2}$ comes from the piecewise definition (5.74) of δ_K, which changes from level 5 to level 6. □

Remark 5.44 (Implementation) The SUPG term influences the velocity-velocity coupling, the pressure (ansatz)—velocity (test) coupling, and the right-hand side for the test function of the velocity. One gets the matrix entries

$$(A_{11})_{ij} = a_{ij} = \sum_{K\in\mathcal{T}^h}\Big[\left(\nu\nabla\boldsymbol{\phi}_j^h,\nabla\boldsymbol{\phi}_i^h\right)_K + \left((\boldsymbol{b}\cdot\nabla)\,\boldsymbol{\phi}_j^h + c\boldsymbol{\phi}_j^h,\boldsymbol{\phi}_i^h\right)_K$$
$$+\left(-\nu\Delta\boldsymbol{\phi}_j^h + (\boldsymbol{b}\cdot\nabla)\,\boldsymbol{\phi}_j^h + c\boldsymbol{\phi}_j^h,\delta_K^v\,(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{\phi}_i^h\right)_K\Big],$$

$i,j = 1,\ldots,3N_v$, and

$$(D)_{ij} = d_{ij} = \sum_{K\in\mathcal{T}^h} -\left[\left(\nabla\cdot\boldsymbol{\phi}_i^h,\psi_j^h\right)_K + \left(\nabla\psi_j^h,\delta_K^v\left(\boldsymbol{b}\cdot\nabla\boldsymbol{\phi}_i^h\right)\right)_K\right],$$

$i = 1,\ldots,3N_v$, $j = 1,\ldots,N_p$. The matrix block of the velocity-velocity coupling has still the diagonal form (5.26). The right-hand side becomes

$$(\underline{f})_i = f_i = \sum_{K\in\mathcal{T}^h}\left[\left(\boldsymbol{f},\boldsymbol{\phi}_i^h\right)_K + \left(\boldsymbol{f},\delta_K^v\,(\boldsymbol{b}\cdot\nabla)\,\boldsymbol{\phi}_i^h\right)_K\right],\quad i = 1,\ldots,3N_v.$$

Altogether, the coupled system has the form

$$\begin{pmatrix} A & D \\ B & 0\end{pmatrix}\begin{pmatrix}\underline{u}\\ \underline{p}\end{pmatrix} = \begin{pmatrix}\underline{f}\\ \underline{0}\end{pmatrix},$$

such that the matrix D has to be stored in addition to A and B.

For the SUPG/PSPG/grad-div method, one has to add all requirements of the individual stabilization terms, see Remarks 4.101 and 4.129. The linear system of equations for this method has the form

$$\begin{pmatrix} A & D \\ B & -C \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{f_p} \end{pmatrix}, \tag{5.75}$$

where the matrix block A is of structure

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{12}^T & A_{22} & A_{23} \\ A_{13}^T & A_{23}^T & A_{33} \end{pmatrix}. \tag{5.76}$$

Altogether, the memory requirements of the SUPG/PSPG/grad-div method are considerably higher compared with the Galerkin method. □

Remark 5.45 (Concluding Remarks on the SUPG/PSPG/grad-div Method)

- In Lube and Rapin (2006), the impact of the polynomial degree of the velocity and pressure finite element spaces on the analytical results is studied. An inf-sup condition similar to (5.52) is proved where the inf-sup constant is independent of the viscosity, the mesh width, and the polynomial degree. Moreover, the polynomial degree enters the definition of the asymptotic optimal stabilization parameters.
- The SUPG/PSPG/grad-div method introduces an artificial nonsymmetric term.
- The SUPG/PSPG/grad-div method introduces a strong coupling of velocity and pressure. There is no physical interpretation for the term $\sum_{K\in\mathcal{T}^h} \delta_K \left\| (\boldsymbol{b}\cdot\nabla)\,\boldsymbol{u}^h + \nabla p^h \right\|_{L^2(K)}^2$.
- Since the residual contains the right-hand side $\boldsymbol{f}$ of the problem, all residual-based stabilizations introduce a modification of the right-hand side. This modification makes the application of these schemes difficult for time-dependent problems.
- The number of matrix blocks which has to be stored and assembled for the SUPG/PSPG/grad-div method is quite large, see Remark 5.44.

□

5.3.3 Other Residual-Based Stabilizations

Remark 5.46 (Neglecting the PSPG Term for Inf-Sup Stable Pairs of Finite Element Spaces) If inf-sup stable pairs of finite element spaces are used, the PSPG term is not necessary since it gives stability if the discrete inf-sup condition (3.51) is not satisfied. The case of inf-sup stable pairs of finite element spaces and the

SUPG/grad-div (sg) stabilization was investigated in Gelhard et al. (2005). It has the form: Find $\left(\boldsymbol{u}^h, p^h\right) \in V^h \times Q^h$ such that

$$A_{\mathrm{sg}}\left(\left(\boldsymbol{u}^h, p^h\right), \left(\boldsymbol{v}^h, q^h\right)\right) = L_{\mathrm{sg}}\left(\boldsymbol{v}^h\right) \quad \forall \ \left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h,$$

with the bilinear form $A_{\mathrm{sg}} : \left(V \times \tilde{Q}\right) \times \left(V \times \tilde{Q}\right) \to \mathbb{R}$ defined by

$$\begin{aligned}
&A_{\mathrm{sg}}\left((\boldsymbol{u}, p), (\boldsymbol{v}, q)\right) \\
&= \nu\left(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}\right) + \left((\boldsymbol{b} \cdot \nabla)\boldsymbol{u} + c\boldsymbol{u}, \boldsymbol{v}\right) - (\nabla \cdot \boldsymbol{v}, p) + (\nabla \cdot \boldsymbol{u}, q) \\
&\quad + \sum_{K \in \mathcal{T}^h} \mu_K \left(\nabla \cdot \boldsymbol{u}, \nabla \cdot \boldsymbol{v}\right)_K \\
&\quad + \sum_{K \in \mathcal{T}^h} \left(-\nu \Delta \boldsymbol{u} + (\boldsymbol{b} \cdot \nabla)\boldsymbol{u} + c\boldsymbol{u} + \nabla p, \delta_K^v (\boldsymbol{b} \cdot \nabla)\boldsymbol{v}\right)_K
\end{aligned}$$

and the linear form $L_{\mathrm{sg}} : V \to \mathbb{R}$ by

$$L_{\mathrm{sg}}(\boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}) + \sum_{K \in \mathcal{T}^h} \left(\boldsymbol{f}, \delta_K^v (\boldsymbol{b} \cdot \nabla)\boldsymbol{v}\right)_K.$$

This scheme is called reduced stabilized scheme in Gelhard et al. (2005). In this paper, error estimates are proved for families of quasi-uniform triangulations and for discrete pressure spaces satisfying $Q^h \subset H^1(\Omega)$. The analysis from Gelhard et al. (2005) was refined in Matthies et al. (2009) such that shape regular grids and discontinuous discrete pressure spaces are covered. An important technical tool in this analysis is the use of the quasi-local interpolation operator from Remark 3.62, which preserves the discrete divergence. However, this operator requires that the polynomial degree of the discrete velocity space is equal or higher than the dimension of the domain Ω. Hence, the most popular pairs of finite element spaces in three dimensions, which use second order velocity, are not covered by this analysis. The way of performing the error analysis of this method is similar to the analysis presented in Sect. 5.3.2 for the SUPG/PSPG/grad-div method. The asymptotic optimal choices of the stabilization parameters, derived from the error analysis, are the same for both methods

$$\delta_K \sim h_K^2, \quad \mu_K \sim 1,$$

see Remark 5.42 for the SUPG/PSPG/grad-div method.

In numerical studies in Gelhard et al. (2005), it turned out that the full SUPG/PSPG/grad-div stabilization and the SUPG/grad-div stabilization give almost identical results for inf-sup stable pairs of finite element spaces. In Matthies et al. (2009), extensive and careful numerical studies on the choice of the stabilization parameters are presented. These studies illustrate in particular that different parameters of the same asymptotic type might lead to errors of different orders of magnitude. These studies also come to the conclusion that the use of only the SUPG stabilization might lead to instabilities and at least one of the other stabilizations, PSPG or grad-div, should be added to the SUPG stabilization. □

Remark 5.47 (Implementation of the SUPG/grad-div Stabilization) Compared with the SUPG/PSPG/grad-div method, there is no block that couples test and ansatz functions for the pressure such that the coupled system has the form

$$\begin{pmatrix} A & D \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{0} \end{pmatrix},$$

where the matrix block A is of structure (5.76). □

5.4 Other Stabilized Finite Element Methods

Remark 5.48 (Motivation) Other stabilizations than residual-based ones try to avoid the main drawbacks of the latter: the introduction of artificial nonsymmetric terms and the strong non-physical coupling of velocity and pressure, see Remark 5.45. In this section, alternative stabilization concepts will be presented and discussed briefly.

A review of stabilization techniques for the Oseen equations can be found in Braack et al. (2007). □

Remark 5.49 (The Continuous Interior Penalty (CIP) Method) The CIP method was already introduced in Douglas and Dupont (1976). Its basic idea consists in increasing the stability of the Galerkin discretization by introducing a least squares control (penalty) of gradient jumps across faces of the mesh cells. Almost three decades later, this approach was applied, e.g., to convection-dominated convection-diffusion equations in Burman and Hansbo (2004) and to the Oseen equations in Burman et al. (2006).

For the Oseen equations, the CIP stabilization, as presented in Burman et al. (2006), reads as follows: Find $\left(\boldsymbol{u}^h, p^h\right) \in V^h \times Q^h$ such that

$$\begin{aligned} &\left(\nu \nabla \boldsymbol{u}^h, \nabla \boldsymbol{v}^h\right) + \left((\boldsymbol{b}\cdot\nabla)\,\boldsymbol{u}^h + c\boldsymbol{u}^h, \boldsymbol{v}^h\right) - \left(\nabla\cdot\boldsymbol{v}^h, p^h\right) + \left(\nabla\cdot\boldsymbol{u}^h, q^h\right) \\ &\quad + \sum_{K\in\mathcal{T}^h} \delta_0 \delta\left(\mathrm{Re}_K\right) h_K^2 \\ &\quad \times \sum_{E\in\partial K} \int_E \|\boldsymbol{b}\cdot\boldsymbol{n}_E\|_{L^\infty(E)} \left[\left|\nabla \boldsymbol{u}^h \boldsymbol{n}_{\partial K}\right|\right]_E \cdot \left[\left|\nabla \boldsymbol{v}^h \boldsymbol{n}_{\partial K}\right|\right]_E \, ds \end{aligned}$$

$$+ \sum_{K\in\mathcal{T}^h} \delta_0 \delta(\mathrm{Re}_K) \|\boldsymbol{b}\|_{L^\infty(K)} h_K^2 \sum_{E\in\partial K} \int_E [|\nabla\cdot\boldsymbol{u}^h|]_E [|\nabla\cdot\boldsymbol{v}^h|]_E \, d\boldsymbol{s}$$

$$+ \sum_{K\in\mathcal{T}^h} \delta_0 \delta(\mathrm{Re}_K) \frac{h_K^2}{\|\boldsymbol{b}\|_{L^\infty(K)}} \sum_{E\in\partial K} \int_E [|\nabla p^h|]_E \cdot [|\nabla q^h|]_E \, d\boldsymbol{s}$$

$$= (\boldsymbol{f}, \boldsymbol{v}^h) \quad \forall\ (\boldsymbol{v}^h, q^h) \in V^h \times Q^h, \tag{5.77}$$

with

$$\mathrm{Re}_K = \frac{\|\boldsymbol{b}\|_{L^\infty(K)} h_K}{\nu}, \quad \delta(\mathrm{Re}_K) = \min\{1, \mathrm{Re}_K\}.$$

The first jump term in (5.77) acts as stabilization of dominating convection, the second jump term gives some control on the violation of the divergence constraint, and the last jump term stabilizes the violation of the discrete inf-sup condition. The stabilization term in (5.77) is symmetric and the right-hand side of the discrete problem is not affected by the stabilization.

The numerical analysis for the case $V^h/Q^h = P_k/P_k$ or $V^h/Q^h = Q_k/Q_k$ was performed in Burman et al. (2006) for weakly imposed Dirichlet boundary conditions. In this approach, the finite element spaces are defined with natural boundary conditions and additional terms defined on the boundary are introduced into the bilinear form which, e.g., penalize the violation of a Dirichlet boundary condition. For large penalty factors, one gets a good approximation of these boundary conditions. Error estimates were derived in a norm which contains contributions from the stabilization terms.

In Braack et al. (2007), a modification of the CIP method is presented which is formulated in terms of sums over the faces

$$(\nu\nabla\boldsymbol{u}^h, \nabla\boldsymbol{v}^h) + ((\boldsymbol{b}\cdot\nabla)\boldsymbol{u}^h + c\boldsymbol{u}^h, \boldsymbol{v}^h) - (\nabla\cdot\boldsymbol{v}^h, p^h) + (\nabla\cdot\boldsymbol{u}^h, q^h)$$

$$+ \sum_{E\in\mathcal{E}^h} \int_E \delta_E^{\boldsymbol{u}} [|\nabla\boldsymbol{u}^h\boldsymbol{n}_{\partial K}|]_E \cdot [|\nabla\boldsymbol{v}^h\boldsymbol{n}_{\partial K}|]_E \, d\boldsymbol{s}$$

$$+ \sum_{E\in\mathcal{E}^h} \int_E \delta_E^{\mathrm{div}} [|\nabla\cdot\boldsymbol{u}^h|]_E [|\nabla\cdot\boldsymbol{v}^h|]_E \, d\boldsymbol{s}$$

$$+ \sum_{E\in\mathcal{E}^h} \int_E \delta_E^{p} [|\nabla p^h\cdot\boldsymbol{n}_E|]_E \cdot [|\nabla q^h\cdot\boldsymbol{n}_E|]_E \, d\boldsymbol{s}$$

$$= (\boldsymbol{f}, \boldsymbol{v}^h) \quad \forall\ (\boldsymbol{v}^h, q^h) \in V^h \times Q^h, \tag{5.78}$$

where the jumps of the gradient of the pressure are replaced by the jumps of the normal derivatives of the pressure. In this paper, even concrete parameters depending on the polynomial degree k are proposed

$$
\begin{aligned}
\delta_E^{\boldsymbol{u}} &= \|\boldsymbol{b}\cdot\boldsymbol{n}_E\|_{L^\infty(E)}\,\frac{h_E^2}{k^\alpha},\\
\delta_E^{\mathrm{div}} &= \|\boldsymbol{b}\|_{L^\infty(E)}\,\frac{h_E^2}{k^\alpha},\\
\delta_E^{p} &= \min\{1,\mathrm{Re}_E\}\,\frac{h_E^2}{\|\boldsymbol{b}\|_{L^\infty(E)}\,k^\alpha},\quad \mathrm{Re}_E = \frac{\|\boldsymbol{b}\|_{L^\infty(E)}\,h_E}{\nu\alpha^{1/2}},
\end{aligned}
\tag{5.79}
$$

with $\alpha = 7/2$.

The CIP stabilization applied to the Crouzeix–Raviart pair of finite element spaces P_1^{nc}/P_0 of the Oseen equations was analyzed in Burman and Hansbo (2006). □

Remark 5.50 (Implementation of the CIP Method) The appearance of jumps in the bilinear form in (5.77) couples degrees of freedom that are not coupled in the standard Galerkin discretization and in the velocity-velocity matrix of the residual-based stabilizations. Consider for example P_1 finite elements, two simplices with the common face E, and the two basis functions which take the value 1 in the vertex opposite to the common face and the value 0 in all other vertices. Then, the common support of these basis functions is E and all integrals in the Galerkin finite element method (5.13) where both basis functions are involved vanish. However, the jumps of the derivatives of these basis functions across E do not vanish and consequently, the jump terms in (5.77) couple these functions. Hence, the sparsity pattern of the matrices of the CIP method is denser than for the Galerkin method and for the velocity-velocity matrix of residual-based stabilizations.

The CIP method adds stabilization terms which couple either velocity test and ansatz functions or pressure test and ansatz functions. The arising coupled system has the form

$$
\begin{pmatrix} A & B^T \\ B & C \end{pmatrix}\begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{0} \end{pmatrix}.
$$

Because of the grad-div term in the CIP method, the block structure of A is as given in (5.76). □

Remark 5.51 (Numerical Studies for the CIP Method) Comprehensive numerical studies of the CIP method applied to the two-dimensional Oseen equations (with $c = 0$) can be found in Umla (2009). In this thesis, the form (5.78) of the CIP method was used and the stabilization parameters (5.79) were scaled with a factor $\delta_0 > 0$. It turned out that the optimal stabilization parameters were generally not given for $\delta_0 = 1$. In some situations, in particular for the finite element pairs P_3/P_3 and Q_3/Q_3, the errors could be reduced by one order of magnitude with the optimal scaling factor compared with $\delta_0 = 1$. Choosing different scaling factors for the different stabilization terms did not lead to notable improvements of the results.

In addition, comparisons with the SUPG/PSPG/grad-div stabilization, see Sect. 5.3.2, were performed in Umla (2009). Generally, the solutions obtained with both stabilizations were of similar accuracy. But sometimes, the pressure was computed somewhat more accurately with the SUPG/PSPG/grad-div stabilization.

In Umla (2009), the CIP stabilization was applied also to pairs of Taylor–Hood finite element spaces. The observed orders of convergence are the same as for the SUPG/PSPG/grad-div method.

The linear systems of equations in the numerical studies of Umla (2009) were solved with the sparse direct solver umfpack, see Remark 9.5. The computing times for the CIP method were considerably higher than for the SUPG/PSPG/grad-div method, often by a factor of about two. The main reason is the denser sparsity pattern of the matrices coming from the CIP stabilization. □

Remark 5.52 (Local Projection Stabilization (LPS) Methods) The main tool of LPS methods is a local projection P^h_{loc} of a finite element space onto another finite element space, which is usually a discontinuous space. With this projection, to so-called fluctuation operator $\left(I - P^h_{\mathrm{loc}}\right)$ is defined. Then, a stabilization of the Galerkin finite element discretization is achieved by adding weighted $L^2(\Omega)$ inner products of fluctuations of quantities of interest.

A short review of the LPS method can be found in Braack and Lube (2009) and a longer presentation in Roos et al. (2008, Chap. IV.4). □

Remark 5.53 (Stabilizing the Violation of the Discrete Inf-Sup Condition with an LPS Method) The LPS method for stabilizing the violation of the discrete inf-sup condition was proposed in Becker and Braack (2001). Applying the general approach of LPS methods, the difference of the gradient of the discrete pressure and a local projection is added to the continuity equation. This idea was already formulated in Codina and Blasco (1997), where a global projection operator was proposed, see Remark 4.111. From the point of view of numerical efficiency, locally computable projections should be preferred.

The numerical analysis of the LPS method requires a number of assumptions, see Becker and Braack (2001), Braack and Lube (2009) for details. The realization of the LPS method as proposed in Becker and Braack (2001) consists in using two triangulations $\mathcal{T}^{2h}$ and $\mathcal{T}^h$ of Ω, where $\mathcal{T}^h$ is a uniform refinement of $\mathcal{T}^{2h}$, which is the so-called two-level method. The mesh cells on the coarse grid are called macro cells $\{M\}$ and each macro cell contains a number of mesh cells on the fine grid. Then, the local $L^2(\Omega)$ projection has the form

$$P_h^{2h} \ : \ L^2(\Omega) \to R^{2h,\mathrm{disc}}, \quad \left(q^h - P_h^{2h} q^h, \psi^h\right) = 0 \quad \forall\ \psi^h \in R^{2h,\mathrm{disc}},$$

where $R^{2h,\text{disc}}$ is a space of discontinuous finite element functions on the coarser grid, e.g., $R^{2h,\text{disc}} = Q_1^{\text{disc}}$ if $V^h = Q_2$. Thus, the LPS term has the form

$$\begin{aligned} S^h_{\text{lps}}(p,q) &= \left(\delta^{1/2}_{\text{lps},p}\left(I - P_h^{2h}\right)\nabla p^h, \delta^{1/2}_{\text{lps},p}\left(I - P_h^{2h}\right)\nabla q^h\right) \\ &= \left(\delta^{1/2}_{\text{lps},p}\kappa_p^h\left(\nabla p^h\right), \delta^{1/2}_{\text{lps},p}\kappa_p^h\left(\nabla q^h\right)\right) \end{aligned} \tag{5.80}$$

for all $p^h, q^h \in Q^h \cap H^1(\Omega)$. Then, the numerical analysis for the LPS method is performed in the norm

$$\|(\boldsymbol{v},q)\|_{\text{lps}} = \left(\|\nabla \boldsymbol{v}\|^2_{L^2(\Omega)} + \|p\|^2_{L^2(\Omega)} + S^h_{\text{lps}}(p,p)\right)^{1/2}.$$

The analysis shows that in the case of using the same finite element spaces for velocity and pressure on the fine grid, then $\delta_{\text{lps},p} = \mathcal{O}\left(h_M^2\right)$ is the asymptotic correct choice, where h_M is the diameter of a macro cell M.

A variant of the LPS method consists in using instead the fluctuation of the gradients of the pressure as in (5.80) the gradient of the fluctuations

$$S^h_{\text{lps}}(p,q) = \left(\delta^{1/2}_{\text{lps},p}\nabla\left(\kappa_p^h p^h\right), \delta^{1/2}_{\text{lps},p}\nabla\left(\kappa_p^h q^h\right)\right).$$

It is mentioned in Braack and Lube (2009) that this method remains optimal for the Stokes equations but its extension is not optimal for the Oseen equations. □

Remark 5.54 (LPS Schemes) An LPS scheme for the Oseen equations was studied in Braack and Burman (2006). In this paper, a numerical analysis for low order discretizations was presented. Concretely, the so-called two-level method was considered with the pairs Q_k/Q_k, $k \in \{1,2\}$, for velocity and pressure and a local projection which maps onto the space Q_{k-1} on the next coarser grid. In Matthies et al. (2007), an LPS approach was proposed which is based on enrichment and uses only one mesh, the so-called one-level method. A unified finite element error analysis for the one-level and two-level approach and for the case of using the same finite spaces for velocity and pressure is presented in this paper. The case of inf-sup stable finite element spaces is studied in Lube et al. (2008), Matthies and Tobiska (2015).

Using the general approach for constructing LPS schemes leads, see Remark 5.52, to the stabilization term

$$\begin{aligned} S^h_{\text{lps}}\left((\boldsymbol{u},p),(\boldsymbol{v},q)\right) = &\left(\delta_{\text{lps},\boldsymbol{u}}\kappa_{\boldsymbol{u}}^h\left((\boldsymbol{b}\cdot\nabla)\,\boldsymbol{u}\right), \kappa_{\boldsymbol{u}}^h\left((\boldsymbol{b}\cdot\nabla)\,\boldsymbol{v}\right)\right) \\ &+ \left(\delta_{\text{lps,div}}\kappa_{\text{div}}^h\left(\nabla\cdot\boldsymbol{u}\right), \kappa_{\text{div}}^h\left(\nabla\cdot\boldsymbol{v}\right)\right) \\ &+ \left(\delta_{\text{lps},p}\kappa_p^h\left(\nabla p^h\right), \kappa_p^h\left(\nabla q^h\right)\right). \end{aligned}$$

For practical reasons, all local projections are defined principally in the same way, either with the one-level or the two-level method. A difference might be the actual polynomial order of the projection space. It can be seen that the stabilization term is symmetric and the stabilization does not affect the right-hand side of the equations.

The finite element error analysis is performed in the mesh-dependent norm

$$\|(\boldsymbol{v},q)\|_{\mathrm{lps}} = \Big(\|\nabla \boldsymbol{v}\|_{L^2(\Omega)}^2 + \left\|c^{1/2}\boldsymbol{v}\right\|_{L^2(\Omega)}^2 + \nu \|p\|_{L^2(\Omega)}^2 + \left\|c^{1/2}p\right\|_{L^2(\Omega)}^2 + S_{\mathrm{lps}}^h\left((\boldsymbol{v},p),(\boldsymbol{v},p)\right)\Big)^{1/2}.$$

The weights of the pressure terms are the same as in the norm $\|\cdot\|_{\mathrm{spg},p}$ used in the SUPG/PSPG/grad-div method, see (5.48). The analysis relies on the construction of an interpolation operator between the finite element spaces and the projection spaces with certain orthogonality properties. A discrete inf-sup condition for the complete bilinear form can be proved. In the case of using the same finite element spaces for velocity and pressure, both of order k, the optimal asymptotic choices of stabilization parameters are

$$\delta_{\mathrm{lps},\boldsymbol{u}} = \mathcal{O}\left(\frac{h_M}{\|\boldsymbol{b}\|_{W^{k,\infty}(M)}}\right), \quad \delta_{\mathrm{lps,div}} = \mathcal{O}\left(h_M\right), \quad \delta_{\mathrm{lps},p} = \mathcal{O}\left(h_M\right),$$

where h_M is the diameter of a macro cell M, see Remark 5.53. With these parameters, one obtains the same order of convergence as for the SUPG/PSPG/grad-div method. However, one should note that only some terms in $\|\cdot\|_{\mathrm{lps}}$ and $\|\cdot\|_{\mathrm{spg},p}$ from (5.48) are the same. □

Remark 5.55 (Two-level vs. One-level LPS Methods) A discussion of possible choices of finite element spaces for two-level and one-level LPS schemes can be found in Matthies et al. (2007). For a two-level LPS scheme, an example is already given in Remark 5.53. One-level LPS schemes rely on enriching a finite element space of order k with mesh cell bubble functions of a sufficiently large polynomial order. Then, the projection space can be chosen to consists of discontinuous functions of order $k-1$.

The use of a two-level method requires a coarse grid, which is not always available in application. In addition, the matrices possess extended stencils, e.g., compared with the SUPG/PSPG/grad-div method. With the one-level method, one obtains a more compact stencil than with the two-level scheme. But on the same grid, the number of degrees of freedom is larger for the one-level method than for other discretizations, like the SUPG/PSPG/grad-div method or the two-level LPS method. □

Remark 5.56 (Implementation of LPS Methods) The two-level LPS method leads to an extended sparsity pattern of the matrices compared with the residual-based stabilizations. All degrees of freedom on mesh cells that belong to the same macro cell are coupled. For the one-level LPS method, the sparsity of the matrices is not extended by connecting degrees of freedom which are usually not connected. However, the enrichment of the spaces leads to more degrees of freedom on the same grid compared with the Galerkin method or other stabilized discretizations.

□

Remark 5.57 (Different Definition of the Fluctuations) Instead of local projections, local approximation operators or quasi interpolants can be used for defining the fluctuations. A method of this type, called higher order term-by-term stabilization method, was proposed and analyzed in Chacón Rebollo et al. (2013). This method has usually a more compact matrix stencil than LPS methods while retaining the same accuracy. □

Remark 5.58 (Upwind Discretizations for Lowest Order Non-conforming Finite Elements) In the case of non-conforming discretizations, the SUPG method for stabilizing dominating convection looses the advantage of not enlarging the stencils of the matrices, because a term which includes the jumps of the functions across the faces is needed. This effect was already observed for the Crouzeix–Raviart finite element P_1^{nc} applied in the discretization of scalar convection-diffusion equations in John et al. (1997). For this reason, the SUPG method is not used for non-conforming finite elements and dominating convection is stabilized in different ways. Most popular are upwind methods, which combine the finite element method with ideas from finite volume methods. Finite element methods of upwind type were proposed the first time, for scalar convection-diffusion equations, in Ohmori and Ushijima (1984).

Consider a scalar non-conforming space of lowest order, i.e., the Crouzeix–Raviart element P_1^{nc}, see Example B.43, or the Rannacher–Turek element Q_1^{rot}, see Example B.53, and a partition of Ω into mesh cells. For P_1^{nc}, the mesh cells are triangles or tetrahedra and for Q_1^{rot} these are quadrilaterals or hexahedra. Let K_1 and K_2 be two neighboring mesh cells with the common face E_i. Then, exactly one standard basis function $\phi_i(\boldsymbol{x})$ of P_1^{nc} or Q_1^{rot} can be assigned to this face. If $\boldsymbol{m}_{E,i}$ is the barycenter of E_i, then $\phi_i(\boldsymbol{m}_{E,i}) = 1$ for P_1^{nc} and the point value oriented Q_1^{rot} element. In the first step of defining upwind methods, the domain is decomposed into so-called dual domains $\{\omega_{\mathrm{dual},i}\}$. A subdomain $\omega_{\mathrm{dual},i}$ is constructed by taking the barycenters S_{K_1} and S_{K_2} of K_1 and K_2, respectively, and connecting them with the vertices of E_i. The union of the two subdomains obtained in this way is $\omega_{\mathrm{dual},i}$, see Fig. 5.12 for the case of a triangular mesh. For faces on the boundary Γ of Ω, the dual domain consists just of one subdomain.

Let Λ_i be the set of all indices $k \neq i$ for which the faces E_k and E_i belong to the same mesh cell. For instance, these are four indices for interior edges in triangular meshes and ten indices for interior faces in hexahedral meshes. Denote the face of $\omega_{\mathrm{dual},i}$ in between E_i and E_k by Γ_{ik} and the corresponding outer unit normal vector by $\boldsymbol{n}_{ik}$, see again Fig. 5.12 for an illustration.

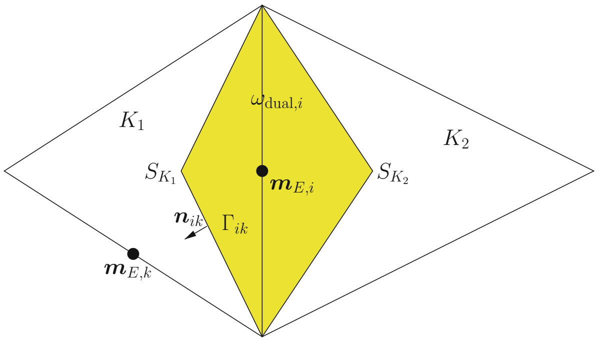

Fig. 5.12 Dual domain for the upwind discretization

Next, an operator L^h is introduced which maps a continuous function $v(\boldsymbol{x})$ onto a function $L^h v(\boldsymbol{x})$ which is constant in each dual domain

$$L^h v(\boldsymbol{x}) = v(\boldsymbol{m}_{E,i}) \quad \forall\, \boldsymbol{x} \in \omega_{\mathrm{dual},i}.$$

To incorporate information about the direction of the convection into the scheme appropriately, the convective term will be approximated with this operator and with integrals on the boundaries of the dual subdomains. One obtains, using the product rule, approximating by applying the operator L^h, utilizing that the approximation is constant in $\omega_{\mathrm{dual},i}$, and applying integration by parts

$$\begin{aligned}
(\boldsymbol{b}\cdot\nabla u, v) &= \sum_{E_i\in\overline{\mathcal{E}}^h} (\boldsymbol{b}\cdot\nabla u, v)_{\omega_{\mathrm{dual},i}} \\
&= \sum_{E_i\in\overline{\mathcal{E}}^h} \Big[(\nabla\cdot(u\boldsymbol{b}), v)_{\omega_{\mathrm{dual},i}} - (\nabla\cdot\boldsymbol{b}, uv)_{\omega_{\mathrm{dual},i}}\Big] \\
&\approx \sum_{E_i\in\overline{\mathcal{E}}^h} \Big[(\nabla\cdot(u\boldsymbol{b}), L^h v)_{\omega_{\mathrm{dual},i}} - (\nabla\cdot\boldsymbol{b}, L^h(uv))_{\omega_{\mathrm{dual},i}}\Big] \\
&= \sum_{E_i\in\overline{\mathcal{E}}^h} \Big((\nabla\cdot(u\boldsymbol{b}), 1)_{\omega_{\mathrm{dual},i}} - u(\boldsymbol{m}_{E,i})(\nabla\cdot\boldsymbol{b}, 1)_{\omega_{\mathrm{dual},i}}\Big)\, v(\boldsymbol{m}_{E,i}) \\
&= \sum_{E_i\in\overline{\mathcal{E}}^h} \sum_{k\in\Lambda_i} \Big((u\boldsymbol{b}\cdot\boldsymbol{n}_{ik}, 1)_{\Gamma_{ik}} - u(\boldsymbol{m}_{E,i})\,(\boldsymbol{b}\cdot\boldsymbol{n}_{ik}, 1)_{\Gamma_{ik}}\Big)\, v(\boldsymbol{m}_{E,i}) \\
&= \sum_{E_i\in\overline{\mathcal{E}}^h} \sum_{k\in\Lambda_i} (\boldsymbol{b}\cdot\boldsymbol{n}_{ik}, u - u(\boldsymbol{m}_{E,i}))_{\Gamma_{ik}}\, v(\boldsymbol{m}_{E,i}).
\end{aligned} \tag{5.81}$$

In upwind discretizations, u will be replaced by a convex combination of the form

$$u \approx \lambda_{ik}(\boldsymbol{b})u(\boldsymbol{m}_{E,i}) + (1 - \lambda_{ik}(\boldsymbol{b}))\, u(\boldsymbol{m}_{E,k}). \tag{5.82}$$

The general upwind idea consists in taking the direction of the convection into account in the discretization. Information which comes from where the convection is coming from (upwind) is weighted stronger than information from where the convection is going to (downwind). In this way, an appropriate transport of information is incorporated in the discretization. The boundary integral $(\boldsymbol{b} \cdot \boldsymbol{n}_{ik}, 1)_{\Gamma_{ik}}$ describes the convective flux across Γ_{ik}. The direction of the flux is determined with the sign of $(\boldsymbol{b} \cdot \boldsymbol{n}_{ik}, 1)_{\Gamma_{ik}}$. If $(\boldsymbol{b} \cdot \boldsymbol{n}_{ik}, 1)_{\Gamma_{ik}} < 0$, then the normal and the convection possess different signs, i.e., the convective flux is directed into $\omega_{\mathrm{dual},i}$. That means, the flux occurs from face E_k to face E_i. For this reason, one chooses $\lambda_{ik} \in [0, 1/2)$ in (5.82) such that the impact of $u(\boldsymbol{m}_{E,k})$ in the discretization of the convective term in the node $\boldsymbol{m}_{E,i}$ is stronger than the impact of $u(\boldsymbol{m}_{E,i})$. In the case $(\boldsymbol{b} \cdot \boldsymbol{n}_{ik}, 1)_{\Gamma_{ik}} > 0$, one chooses with analogous considerations $\lambda_{ik} \in (1/2, 1]$. Thus, one obtains by inserting (5.82) in (5.81) the following discretization of $(\boldsymbol{b} \cdot \nabla u, v)$

$$\begin{aligned}
\left(\boldsymbol{b} \cdot \nabla u^h, v^h\right) &\approx n^h_{\mathrm{upw}}\left(\boldsymbol{b}, u^h, v^h\right) \\
&= \sum_{E_i \in \overline{\mathcal{E}}^h} \sum_{k \in \Lambda_i} \left(\boldsymbol{b} \cdot \boldsymbol{n}_{ik}, (1 - \lambda_{ik}(\boldsymbol{b}))(u^h(\boldsymbol{m}_{E,k}) - u^h(\boldsymbol{m}_{E,i})\right)_{\Gamma_{ik}} v^h(\boldsymbol{m}_{E,i}).
\end{aligned}$$

Defining

$$\lambda_{ik}(\boldsymbol{b}) = \Phi(t) \quad \text{with} \quad t = \frac{1}{2\nu}\left(\boldsymbol{b} \cdot \boldsymbol{n}_{ik}, 1\right)_{\Gamma_{ik}},$$

then the function $\Phi(t)$ has to satisfy the following conditions

i) $\Phi(t) = 1 - \Phi(-t)$ for all $t > 0$ and $0 \le \Phi(t) \le 1$ for all $t \in \mathbb{R}$,
ii) $t\left(\Phi(t) - \frac{1}{2}\right) \ge 0$ for all $t \in \mathbb{R}$,
iii) $g(t) = t\Phi(t)$ is Lipschitz continuous in $\mathbb{R}$,

see Roos et al. (2008, Chap. III.3.1). With the ansatz

$$\Phi(t) = \frac{1}{2} + \operatorname{sgn}(t)\Psi(|t|),$$

one obtains

$$1 - \Phi(-t) = \frac{1}{2} - \operatorname{sgn}(-t)\Psi(|-t|) = \frac{1}{2} + \operatorname{sgn}(t)\Psi(|t|) = \Phi(t),$$

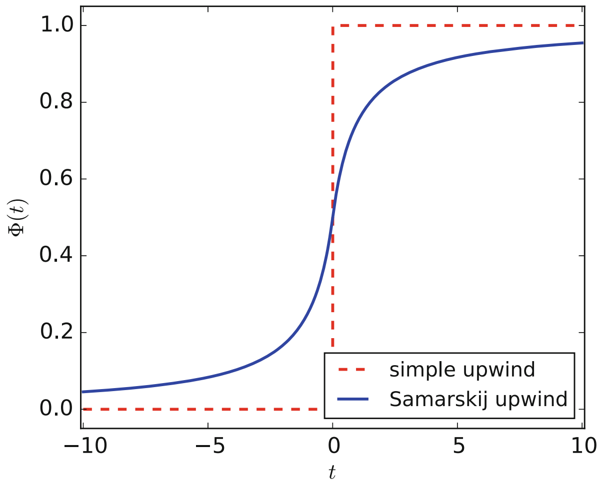

Fig. 5.13 Upwind functions

such that the first part of condition i) is satisfied. Some upwind methods which are often used are the simple or standard upwind and the Samarskij upwind given by

$$\Psi(t) = \frac{1}{2} \quad \text{and} \quad \Psi(t) = \frac{1}{2}\frac{t}{t+1},$$

respectively, see Fig. 5.13.

For the Oseen equations, the upwind methodology is applied to each component of the term $\left((\boldsymbol{b}\cdot\nabla)\,\boldsymbol{u}^h, \boldsymbol{v}^h\right)$.

The convergence of the upwind method applied to the Navier–Stokes equations and error estimates were proved in Schieweck and Tobiska (1989), Schieweck and Tobiska (1996). An overview of the results can be also found in (Roos et al. 2008, Chap. IV.2). Based on the construction of a divergence-preserving interpolation operator, the error estimate

$$\left(\sum_{K\in\mathcal{T}^h} \left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(K)}^2\right)^{1/2} + \left\|p-p^h\right\|_{L^2(\Omega)} \le Ch \tag{5.83}$$

was proved for the stationary Navier–Stokes equations in Schieweck and Tobiska (1996). The constant in (5.83) depends on inverse powers of ν, $\|\boldsymbol{u}\|_{H^2(\Omega)}$, and $\|p\|_{H^1(\Omega)}$. □

Example 5.59 (Numerical Results for an Upwind Discretization) The same setup as described in Example 5.18 is considered. Results obtained with the Samarskij upwind discretization applied for the Crouzeix–Raviart pair of finite element spaces P_1^{nc}/P_0 are presented. The simulations were performed on the irregular triangular grid depicted in Fig. 4.2.

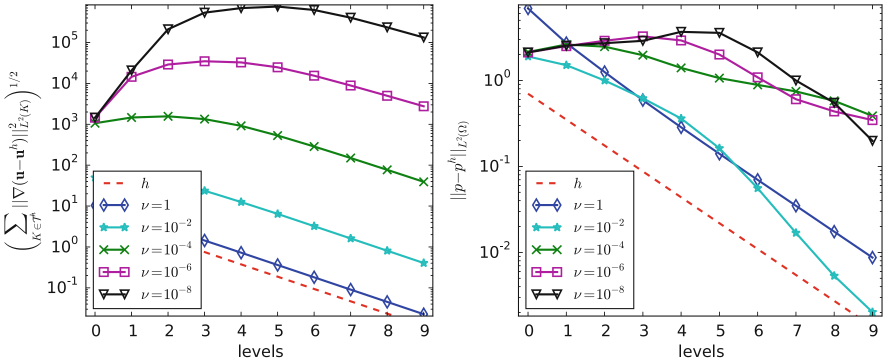

Fig. 5.14 Example 5.59. Samarskij upwind discretization with P_1^{nc}/P_0, convergence of the errors $\|\boldsymbol{u}-\boldsymbol{u}^h\|_{V^h} = \left(\sum_{K\in\mathcal{T}^h} \|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\|_{L^2(K)}^2\right)^{1/2}$ and $\|p-p^h\|_{L^2(\Omega)}$ for $c=0$ and different values of ν

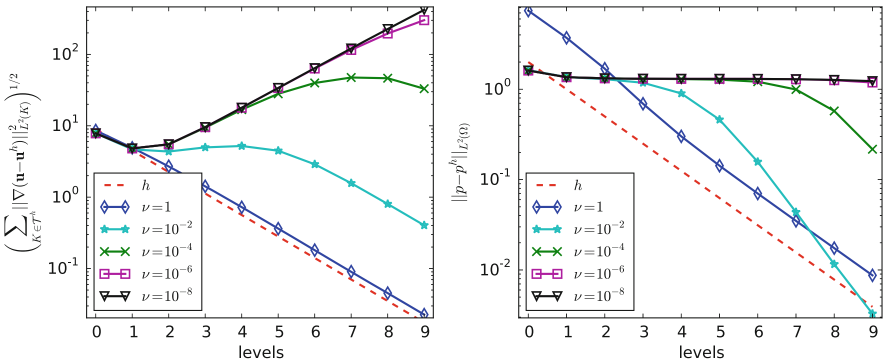

Fig. 5.15 Example 5.59. Samarskij upwind discretization with P_1^{nc}/P_0, convergence of the errors $\|\boldsymbol{u}-\boldsymbol{u}^h\|_{V^h} = \left(\sum_{K\in\mathcal{T}^h} \|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\|_{L^2(K)}^2\right)^{1/2}$ and $\|p-p^h\|_{L^2(\Omega)}$ for $c=100$ and different values of ν

Results for different values of ν and $c=0$ are presented in Fig. 5.14. It can be seen clearly that the velocity error depends on ν^{-1}. Likewise, a dependency of the pressure error on the inverse of ν can be observed, which is however much smaller. The velocity error converges of first order for all values of ν, as predicted in (5.83). A first order convergence of the pressure error can be seen only for $\nu=1$. In all other cases, the asymptotic range of convergence does not seem to be reached yet.

The results for $c=100$ are depicted in Fig. 5.15. For small ν and coarse grids, the velocity error increases if the grid is refined and the pressure error is more or less constant. Only if the grid becomes sufficiently fine, where fine has to be understood with respect to ν, both errors start to decrease. For $\nu=1$, a first order convergence of both errors can be observed.

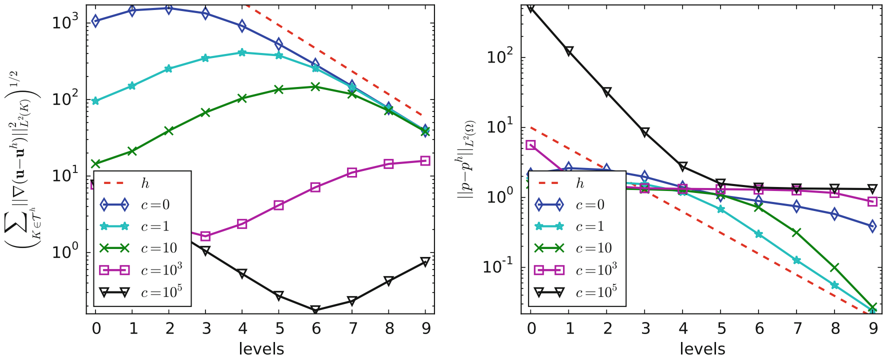

Fig. 5.16 Example 5.59. Samarskij upwind discretization with P_1^{nc}/P_0, convergence of the errors $\|\boldsymbol{u} - \boldsymbol{u}^h\|_{V^h} = \left(\sum_{K\in\mathcal{T}^h} \|\nabla(\boldsymbol{u} - \boldsymbol{u}^h)\|_{L^2(K)}^2\right)^{1/2}$ and $\|p - p^h\|_{L^2(\Omega)}$ for $\nu = 10^{-4}$ and different values of c

Results for $\nu = 10^{-4}$ and different values of c are presented in Fig. 5.16. One can see that large values of c (which correspond to small time steps) lead to small velocity errors on coarse grids but also to large pressure errors.

Evaluating the results presented in Figs. 5.15 and 5.16 from the point of view that c represents the inverse of the time step, one can draw the following conclusions. If the length of the time step is constant, then one needs for small values of ν fine grids for computing reasonably accurate results. For fixed viscosity, small time steps improve the velocity error on coarse grids but not on finer grids. The pressure error stays large for small time steps until the grid becomes sufficiently fine. □

Remark 5.60 (On the Matrices of Upwind Methods for Non-conforming Finite Element Spaces of Lowest Order) In the upwind method, one obtains the same structure of the coupled system as for the Galerkin method, see (5.25) and (5.26). In addition, the maximal number of entries in each row and column of the matrices is independent of the actual meshes since the degrees of freedom can be assigned to the faces and the number of neighboring faces does not depend on the mesh. For instance, for the Crouzeix–Raviart element P_1^{nc}/P_0 there are at most five entries in each row of A_{11} in two dimensions and seven entries in three dimensions. □

Chapter 6
The Steady-State Navier–Stokes Equations

Remark 6.1 (The Steady-State Navier–Stokes Equations) The steady-state or stationary Navier–Stokes equations describe steady-state flows. Such flow fields can be expected in practice if:

- all data of the Navier–Stokes equations (2.25) do not depend on the time,
- the viscosity ν is sufficiently large, or equivalently, the Reynolds number Re is sufficiently small,

see Remark 2.22.

The Navier–Stokes equations are nonlinear. That means, the second difficulty mentioned in Remark 2.19 has to be addressed. □

6.1 The Continuous Equations

Remark 6.2 (Monograph) A comprehensive presentation of the analysis of the steady-state Navier–Stokes equations, and of the Stokes and Oseen equations as well, can be found in Galdi (2011). □

6.1.1 The Strong Form and the Variational Form

Remark 6.3 (Strong Form of the Steady-State Navier–Stokes Equations) The steady-state Navier–Stokes equations are given by

$$\begin{aligned} -\nu\Delta\boldsymbol{u} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p &= \boldsymbol{f} \text{ in } \Omega, \\ \nabla\cdot\boldsymbol{u} &= 0 \text{ in } \Omega, \end{aligned} \tag{6.1}$$

V. John, *Finite Element Methods for Incompressible Flow Problems*, Springer Series in Computational Mathematics 51, DOI 10.1007/978-3-319-45750-5_6

where $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, is a bounded domain with Lipschitz boundary. As for the Stokes and Oseen equations, the numerical analysis will be presented for the case of homogeneous Dirichlet boundary conditions $\boldsymbol{u} = \boldsymbol{0}$ on Γ. □

Remark 6.4 (Variational Form of the Steady-State Navier–Stokes Equations) For the variational formulation of the steady-state Navier–Stokes equations, the same function spaces $V = H_0^1(\Omega)$ and $Q = L_0^2(\Omega)$ as for the Stokes and Oseen equations can be used. A variational form of (6.1) is as follows: Find $(\boldsymbol{u}, p) \in V \times Q$ such that

$$(\nu \nabla \boldsymbol{u}, \nabla \boldsymbol{v}) + ((\boldsymbol{u} \cdot \nabla)\boldsymbol{u}, \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, p) = \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{V',V},$$
$$-(\nabla \cdot \boldsymbol{u}, q) = 0 \tag{6.2}$$

for all $(\boldsymbol{v}, q) \in V \times Q$. The problem can be formulated equivalently in the following way: Find $(\boldsymbol{u}, p) \in V \times Q$ such that

$$(\nu \nabla \boldsymbol{u}, \nabla \boldsymbol{v}) + ((\boldsymbol{u} \cdot \nabla)\boldsymbol{u}, \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, p) + (\nabla \cdot \boldsymbol{u}, q) = \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{V',V} \tag{6.3}$$

for all $(\boldsymbol{v}, q) \in V \times Q$.

By the Sobolev imbedding $H^1(\Omega) \to L^4(\Omega)$, see (A.21) and (A.22), one has that $\boldsymbol{u}, \boldsymbol{v} \in L^4(\Omega)$. Since $\nabla \boldsymbol{u} \in L^2(\Omega)$, it follows from the generalized Hölder inequality, compare (6.33), that the term $((\boldsymbol{u} \cdot \nabla)\boldsymbol{u}, \boldsymbol{v})$ is well defined if $\boldsymbol{u}, \boldsymbol{v} \in V$. □

Remark 6.5 (The Reduced Problem in V_{div}) There is also an associated problem in the space V_{div} of weakly divergence-free functions: Find $\boldsymbol{u} \in V_{\mathrm{div}}$ such that

$$(\nu \nabla \boldsymbol{u}, \nabla \boldsymbol{v}) + ((\boldsymbol{u} \cdot \nabla)\boldsymbol{u}, \boldsymbol{v}) = \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{V',V} \quad \forall\ \boldsymbol{v} \in V_{\mathrm{div}}. \tag{6.4}$$

Of course, if $(\boldsymbol{u}, p) \in V \times Q$ is a solution of (6.3), $\boldsymbol{u} \in V_{\mathrm{div}}$ and $\boldsymbol{u}$ is a solution of (6.4). The other direction can be proved as for linear saddle point problems, see Sect. 3.1. The bilinear form which couples velocity and pressure is the same as for the Stokes and the Oseen equations and the spaces are the same, too. Thus, given a solution $\boldsymbol{u}$ of (6.4), there exists a unique pressure $p \in Q$ such that $(\boldsymbol{u}, p)$ solves (6.3) since the spaces V and Q satisfy the inf-sup condition (3.14), compare Theorem 3.46. Note that in the proof of the inf-sup condition, see the proof of Lemma 3.12, the bilinear form which couples the ansatz and test functions of the velocity space does not play any role. □

6.1.2 The Nonlinear Term

Remark 6.6 (Different Forms of the Convective Term in (6.1)) There are several forms of the convective term of the incompressible Navier–Stokes equations:

$$(\boldsymbol{u} \cdot \nabla)\boldsymbol{u} \qquad : \qquad \text{convective form},$$

$$\nabla \cdot \left(\boldsymbol{u}\boldsymbol{u}^T\right) \quad : \quad \text{divergence form,}$$

$$\nabla \cdot \left(\boldsymbol{u}\boldsymbol{u}^T\right) + \frac{1}{2}\nabla\left(\boldsymbol{u}^T\boldsymbol{u}\right) \quad : \quad \text{divergence form with modified pressure,}$$

$$(\nabla \times \boldsymbol{u}) \times \boldsymbol{u} \quad : \quad \text{rotational form (with modified pressure).}$$

The equivalence of these forms will be established now.

Let $\boldsymbol{u} = (u_1, u_2, u_3)^T$ be weakly differentiable and weakly divergence-free, i.e., $\nabla \cdot \boldsymbol{u} = 0$ almost everywhere.

- *Convective form and divergence form.* The convective form and the divergence form are equivalent, since one gets with the product rule

$$\begin{aligned}
&\nabla \cdot \left(\boldsymbol{u}\boldsymbol{u}^T\right) \\
&= \nabla \cdot \begin{pmatrix} u_1u_1 & u_1u_2 & u_1u_3 \\ u_2u_1 & u_2u_2 & u_2u_3 \\ u_3u_1 & u_3u_2 & u_3u_3 \end{pmatrix} = \begin{pmatrix} \partial_x(u_1u_1) + \partial_y(u_1u_2) + \partial_z(u_1u_3) \\ \partial_x(u_2u_1) + \partial_y(u_2u_2) + \partial_z(u_2u_3) \\ \partial_x(u_3u_1) + \partial_y(u_3u_2) + \partial_z(u_3u_3) \end{pmatrix} \\
&= \begin{pmatrix} (u_1\partial_x + u_2\partial_y + u_3\partial_z)u_1 \\ (u_1\partial_x + u_2\partial_y + u_3\partial_z)u_2 \\ (u_1\partial_x + u_2\partial_y + u_3\partial_z)u_3 \end{pmatrix} + \begin{pmatrix} u_1(\partial_x(u_1) + \partial_y(u_2) + \partial_z(u_3)) \\ u_2(\partial_x(u_1) + \partial_y(u_2) + \partial_z(u_3)) \\ u_3(\partial_x(u_1) + \partial_y(u_2) + \partial_z(u_3)) \end{pmatrix} \\
&= (\boldsymbol{u} \cdot \nabla)\boldsymbol{u} + (\nabla \cdot \boldsymbol{u})\boldsymbol{u} = (\boldsymbol{u} \cdot \nabla)\boldsymbol{u}. \qquad (6.5)
\end{aligned}$$

- *Convective form and divergence form with modified pressure.* From (6.5), one obtains

$$\begin{aligned}
(\boldsymbol{u} \cdot \nabla)\boldsymbol{u} + \nabla p &= \nabla \cdot \left(\boldsymbol{u}\boldsymbol{u}^T\right) + \frac{1}{2}\nabla\left(\boldsymbol{u}^T\boldsymbol{u}\right) - \frac{1}{2}\nabla\left(\boldsymbol{u}^T\boldsymbol{u}\right) + \nabla p \\
&= \nabla \cdot \left(\boldsymbol{u}\boldsymbol{u}^T\right) + \frac{1}{2}\left(\boldsymbol{u}^T\boldsymbol{u}\right) + \nabla p_{\text{mod}}
\end{aligned}$$

with

$$p_{\text{mod}} = p - \frac{1}{2}\left(\boldsymbol{u}^T\boldsymbol{u}\right).$$

- *Convective form and rotational form.* The rotational form goes also along with a redefinition of the pressure. In contrast to the equivalence of the convective form and the divergence form, it is not required that $\boldsymbol{u}$ is divergence-free. It holds

$$\begin{aligned}
(\nabla \times \boldsymbol{u}) \times \boldsymbol{u} + \frac{1}{2}\nabla\left(\boldsymbol{u}^T\boldsymbol{u}\right) &= \begin{pmatrix} \partial_y u_3 - \partial_z u_2 \\ \partial_z u_1 - \partial_x u_3 \\ \partial_x u_2 - \partial_y u_1 \end{pmatrix} \times \boldsymbol{u} + \frac{1}{2}\nabla\left(u_1^2 + u_2^2 + u_3^2\right) \\
&= \begin{pmatrix} u_3\partial_z u_1 - u_3\partial_x u_3 - u_2\partial_x u_2 + u_2\partial_y u_1 \\ u_1\partial_x u_2 - u_1\partial_y u_1 - u_3\partial_y u_3 + u_3\partial_z u_2 \\ u_2\partial_y u_3 - u_2\partial_z u_2 - u_1\partial_z u_1 + u_1\partial_x u_3 \end{pmatrix} \\
&\quad + \frac{1}{2}\begin{pmatrix} 2u_1\partial_x u_1 + 2u_2\partial_x u_2 + 2u_3\partial_x u_3 \\ 2u_1\partial_y u_1 + 2u_2\partial_y u_2 + 2u_3\partial_y u_3 \\ 2u_1\partial_z u_1 + 2u_2\partial_z u_2 + 2u_3\partial_z u_3 \end{pmatrix} \\
&= \begin{pmatrix} u_1\partial_x u_1 + u_2\partial_y u_1 + u_3\partial_z u_1 \\ u_1\partial_x u_2 + u_2\partial_y u_2 + u_3\partial_z u_2 \\ u_1\partial_x u_3 + u_2\partial_y u_3 + u_3\partial_z u_3 \end{pmatrix} \\
&= (\boldsymbol{u} \cdot \nabla)\boldsymbol{u}. \qquad (6.6)
\end{aligned}$$

The term with the gradient is used to define a new pressure, the so-called Bernoulli pressure

$$p_{\text{Bern}} = p + \frac{1}{2}\boldsymbol{u}^T\boldsymbol{u}. \qquad (6.7)$$

The derivation of identity (6.6) is also possible on the basis of (3.158).

□

Lemma 6.7 (Basic Properties of the Convective Term)

- *Let $\boldsymbol{v}$ be weakly differentiable, then it is*

$$(\boldsymbol{u} \cdot \nabla)\,\boldsymbol{v} = (\nabla \boldsymbol{v})\,\boldsymbol{u}. \qquad (6.8)$$

- *Let $\boldsymbol{u}$ be weakly differentiable, then it holds*

$$(\boldsymbol{u} \cdot \nabla)\,\boldsymbol{u} = \mathbb{D}\,(\boldsymbol{u})\,\boldsymbol{u} + \frac{1}{2}\,(\nabla \cdot \boldsymbol{u}) \times \boldsymbol{u}. \qquad (6.9)$$

- *The variational form of the convective term is trilinear, i.e., it is linear in each argument.*
- *Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in H^1(\Omega)$, then*

$$((\boldsymbol{u} \cdot \nabla)\,\boldsymbol{v}, \boldsymbol{w}) = \left(\nabla \cdot \left(\boldsymbol{v}\boldsymbol{u}^T\right), \boldsymbol{w}\right) - ((\nabla \cdot \boldsymbol{u})\,\boldsymbol{v}, \boldsymbol{w})\,, \qquad (6.10)$$

$$((\boldsymbol{u} \cdot \nabla)\,\boldsymbol{v}, \boldsymbol{w}) = (\boldsymbol{u}, \nabla\,(\boldsymbol{v} \cdot \boldsymbol{w})) - ((\boldsymbol{u} \cdot \nabla)\,\boldsymbol{w}, \boldsymbol{v})\,, \qquad (6.11)$$

$$((\boldsymbol{u}\cdot\nabla)\,\boldsymbol{v},\boldsymbol{w}) = \left((\nabla\boldsymbol{v})^T\,\boldsymbol{w},\boldsymbol{u}\right), \tag{6.12}$$

$$((\boldsymbol{u}\cdot\nabla)\,\boldsymbol{v},\boldsymbol{w}) = \int_{\partial\Omega}(\boldsymbol{u}\cdot\boldsymbol{n})\,(\boldsymbol{v}\cdot\boldsymbol{w})\;ds - (\nabla\cdot\boldsymbol{u},\boldsymbol{v}\cdot\boldsymbol{w}) - ((\boldsymbol{u}\cdot\nabla)\,\boldsymbol{w},\boldsymbol{v}), \tag{6.13}$$

where $\boldsymbol{n}$ *is the outward pointing unit normal at* $\partial\Omega$*.*

- *Let* $\boldsymbol{u},\boldsymbol{v}\in H^1(\Omega)$*,* $\boldsymbol{w}\in V_{\mathrm{div}}$*, then*

$$((\nabla\times\boldsymbol{u})\times\boldsymbol{v},\boldsymbol{w}) = ((\boldsymbol{v}\cdot\nabla)\,\boldsymbol{u},\boldsymbol{w}) - ((\boldsymbol{w}\cdot\nabla)\,\boldsymbol{u},\boldsymbol{v}). \tag{6.14}$$

Proof

- Relation (6.8) is proved with a direct calculation

$$(\boldsymbol{u}\cdot\nabla)\,\boldsymbol{v} = \begin{pmatrix} u_1\partial_x v_1 + u_2\partial_y v_1 + u_3\partial_z v_1 \\ u_1\partial_x v_2 + u_2\partial_y v_2 + u_3\partial_z v_2 \\ u_1\partial_x v_3 + u_2\partial_y v_3 + u_3\partial_z v_3 \end{pmatrix} = \begin{pmatrix} \partial_x v_1 & \partial_y v_1 & \partial_z v_1 \\ \partial_x v_2 & \partial_y v_2 & \partial_z v_2 \\ \partial_x v_3 & \partial_y v_3 & \partial_z v_3 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = (\nabla\boldsymbol{v})\,\boldsymbol{u}.$$

- With (6.8), one finds that

$$\mathbb{D}\,(\boldsymbol{u})\,\boldsymbol{u} = \frac{(\nabla\boldsymbol{u})\,\boldsymbol{u}}{2} + \frac{\left(\nabla\boldsymbol{u}^T\right)\boldsymbol{u}}{2} = \frac{(\boldsymbol{u}\cdot\nabla)\,\boldsymbol{u}}{2} + \frac{\left(\nabla\boldsymbol{u}^T\right)\boldsymbol{u}}{2}. \tag{6.15}$$

An intermediate result of (6.6) is $\nabla\left(\boldsymbol{u}^T\boldsymbol{u}\right) = 2\left(\nabla\boldsymbol{u}^T\right)\boldsymbol{u}$ (product rule). Then, one gets from (6.6) and (6.15)

$$(\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = (\nabla\times\boldsymbol{u})\times\boldsymbol{u} + \left(\nabla\boldsymbol{u}^T\right)\boldsymbol{u} = (\nabla\times\boldsymbol{u})\times\boldsymbol{u} + 2\mathbb{D}\,(\boldsymbol{u})\,\boldsymbol{u} - (\boldsymbol{u}\cdot\nabla)\boldsymbol{u},$$

which is equivalent to (6.9).

- The trilinearity of the convective term follows by the linearity of differentiation and integration, e.g., for $a,b\in\mathbb{R}$, one obtains

$$\begin{aligned}
&(((a\tilde{\boldsymbol{u}} + b\hat{\boldsymbol{u}})\cdot\nabla)\,\boldsymbol{v},\boldsymbol{w}) \\
&= \int_\Omega \begin{pmatrix} (a\tilde{u} + b\hat{u})_1\partial_x v_1 + (a\tilde{u} + b\hat{u})_2\partial_y v_1 + (a\tilde{u} + \hat{u})_3\partial_z v_1 \\ (a\tilde{u} + b\hat{u})_1\partial_x v_2 + (a\tilde{u} + b\hat{u})_2\partial_y v_2 + (a\tilde{u} + b\hat{u})_3\partial_z v_2 \\ (a\tilde{u} + b\hat{u})_1\partial_x v_3 + (a\tilde{u} + b\hat{u})_2\partial_y v_3 + (a\tilde{u} + b\hat{u})_3\partial_z v_3 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} d\boldsymbol{x} \\
&= \int_\Omega a \begin{pmatrix} \tilde{u}_1\partial_x v_1 + \tilde{u}_2\partial_y v_1 + \tilde{u}_3\partial_z v_1 \\ \tilde{u}_1\partial_x v_2 + \tilde{u}_2\partial_y v_2 + \tilde{u}_3\partial_z v_2 \\ \tilde{u}_1\partial_x v_3 + \tilde{u}_2\partial_y v_3 + \tilde{u}_3\partial_z v_3 \end{pmatrix} d\boldsymbol{x} + \int_\Omega b \begin{pmatrix} \hat{u}_1\partial_x v_1 + \hat{u}_2\partial_y v_1 + \hat{u}_3\partial_z v_1 \\ \hat{u}_1\partial_x v_2 + \hat{u}_2\partial_y v_2 + \hat{u}_3\partial_z v_2 \\ \hat{u}_1\partial_x v_3 + \hat{u}_2\partial_y v_3 + \hat{u}_3\partial_z v_3 \end{pmatrix} d\boldsymbol{x} \\
&= a\,((\tilde{\boldsymbol{u}}\cdot\nabla)\,\boldsymbol{v},\boldsymbol{w}) + b\,((\hat{\boldsymbol{u}}\cdot\nabla)\,\boldsymbol{v},\boldsymbol{w}).
\end{aligned}$$

The calculations for the two other arguments of the convective term are similar.

- For $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in H^1(\Omega)$, the relations (6.10) and (6.11) are directly proved by straightforward calculations like in Remark 6.6. For instance, using the product rule and rearranging terms yields

$$\begin{aligned}
&(\boldsymbol{u}, \nabla (\boldsymbol{v} \cdot \boldsymbol{w})) \\
&= \int_\Omega \boldsymbol{u} \cdot \nabla (v_1 w_1 + v_2 w_2 + v_3 w_3) \; d\boldsymbol{x} \\
&= \int_\Omega \Big[u_1 v_1 \partial_x w_1 + u_1 v_2 \partial_x w_2 + u_1 v_3 \partial_x w_3 + u_2 v_1 \partial_y w_1 + u_1 v_2 \partial_y w_2 \\
&\quad + u_2 v_3 \partial_y w_3 + u_3 v_1 \partial_z w_1 + u_3 v_2 \partial_z w_2 + u_3 v_3 \partial_z w_3 \Big] \\
&\quad + \Big[u_1 w_1 \partial_x v_1 + u_1 w_2 \partial_x v_2 + u_1 w_3 \partial_x v_3 + u_2 w_1 \partial_y v_1 + u_1 w_2 \partial_y v_2 \\
&\quad + u_2 w_3 \partial_y v_3 + u_3 w_1 \partial_z v_1 + u_3 w_2 \partial_z v_2 + u_3 w_3 \partial_z v_3 \Big] \, d\boldsymbol{x} \\
&= ((\boldsymbol{u} \cdot \nabla) \boldsymbol{w}, \boldsymbol{v}) + ((\boldsymbol{u} \cdot \nabla) \boldsymbol{v}, \boldsymbol{w}) .
\end{aligned}$$

  Equation (6.12) can be also proved by using (6.8)

$$((\boldsymbol{u} \cdot \nabla) \boldsymbol{v}, \boldsymbol{w}) = ((\nabla \boldsymbol{v}) \boldsymbol{u}, \boldsymbol{w}) = \left(\boldsymbol{u}, (\nabla \boldsymbol{v})^T \boldsymbol{w}\right) = \left((\nabla \boldsymbol{v})^T \boldsymbol{w}, \boldsymbol{u}\right).$$

  Relation (6.13) is a direct consequence of applying integration by parts for the term $(\nabla \cdot \boldsymbol{u}, \boldsymbol{v} \cdot \boldsymbol{w})$ and (6.11).
- For proving (6.14), one starts with the identity (3.158) which gives

$$((\nabla \times \boldsymbol{u}) \times \boldsymbol{v}, \boldsymbol{w}) = ((\boldsymbol{v} \cdot \nabla) \boldsymbol{u}, \boldsymbol{w}) - (\nabla (\boldsymbol{u} \cdot \boldsymbol{v}), \boldsymbol{w}) + \left((\nabla \boldsymbol{v})^T \boldsymbol{u}, \boldsymbol{w}\right). \tag{6.16}$$

  Integration by parts, using that $\boldsymbol{w}$ vanishes on the boundary and is divergence-free in Ω, yields

$$(\nabla (\boldsymbol{u} \cdot \boldsymbol{v}), \boldsymbol{w}) = \int_\Gamma (\boldsymbol{u} \cdot \boldsymbol{v}) (\boldsymbol{w} \cdot \boldsymbol{n}) \; ds - (\boldsymbol{u} \cdot \boldsymbol{v}, \nabla \cdot \boldsymbol{w}) = 0. \tag{6.17}$$

  With a direct calculation, one obtains

$$\begin{aligned}
\left((\nabla \boldsymbol{v})^T \boldsymbol{u}, \boldsymbol{w}\right) &= \partial_x v_1 u_1 w_1 + \partial_x v_2 u_2 w_1 + \partial_x v_3 u_3 w_1 + \partial_y v_1 u_1 w_2 + \partial_y v_2 u_2 w_2 \\
&\quad + \partial_y v_3 u_3 w_2 + \partial_z v_1 u_1 w_3 + \partial_z v_2 u_2 w_3 + \partial_z v_3 u_3 w_3 \\
&= \left(\nabla \boldsymbol{v}, \boldsymbol{u} \boldsymbol{w}^T\right).
\end{aligned} \tag{6.18}$$

  Integration by parts, observing that the integral on Γ vanishes, and (2.28) gives

$$\left(\nabla \boldsymbol{v}, \boldsymbol{u} \boldsymbol{w}^T\right) = -\left(\boldsymbol{v}, \nabla \cdot \left(\boldsymbol{u} \boldsymbol{w}^T\right)\right) = -\left(\boldsymbol{v}, (\boldsymbol{w} \cdot \nabla) \boldsymbol{u}\right) - \left(\boldsymbol{v}, (\nabla \cdot \boldsymbol{w}) \boldsymbol{u}\right). \tag{6.19}$$

The last term vanishes since $\boldsymbol{w}$ is divergence-free. Inserting (6.17) and (6.19) in (6.16) proves (6.14).

■

Remark 6.8 (Equivalent Variational Forms of the Steady-State Navier–Stokes Equations) Formulation (6.2) utilizes the convective form of the nonlinear term

$$n_{\text{conv}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = ((\boldsymbol{u} \cdot \nabla)\boldsymbol{v}, \boldsymbol{w}).$$

Equivalent variational forms are obtained by using the other forms of the nonlinear term described in Remark 6.6. The divergence form is defined by

$$n_{\text{div}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = n_{\text{conv}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) + \frac{1}{2}\left((\nabla \cdot \boldsymbol{u})\, \boldsymbol{v}, \boldsymbol{w}\right). \tag{6.20}$$

The motivation for changing the factor in front of the divergence term in comparison with (6.5) is that (6.20) will vanish if $\boldsymbol{v} = \boldsymbol{w}$ even if the first argument is not weakly divergence-free, see Lemma 6.10. This property is not true if a different factor than $1/2$ is used. If the rotational form

$$n_{\text{rot}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = ((\nabla \times \boldsymbol{u}) \times \boldsymbol{v}, \boldsymbol{w}) \tag{6.21}$$

is applied, then the momentum equation in (6.2) changes to

$$(\nu \nabla \boldsymbol{u}, \nabla \boldsymbol{v}) + n_{\text{rot}}(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}) - (\nabla \cdot \boldsymbol{v}, p_{\text{Bern}}) = \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{V',V} \quad \forall\ \boldsymbol{v} \in V,$$

where the Bernoulli pressure is defined in (6.7).

Finally, for the variational form of the Navier–Stokes equations, another form of the convective term can be applied. Integration by parts gives for $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in H^1(\Omega)$

$$(\nabla \cdot \boldsymbol{u}, \boldsymbol{v} \cdot \boldsymbol{w}) = \int_\Gamma (\boldsymbol{v} \cdot \boldsymbol{w})(\boldsymbol{u} \cdot \boldsymbol{n})\, ds - (\boldsymbol{u}, \nabla(\boldsymbol{v} \cdot \boldsymbol{w})), \tag{6.22}$$

where $\boldsymbol{n}$ is the outward pointing unit normal vector on Γ. From the Sobolev imbeddings $H^1(\Omega) \to L^4(\Omega)$, see (A.21) and (A.22), it follows that $\boldsymbol{v}, \boldsymbol{w} \in L^4(\Omega)$ and consequently that $\boldsymbol{v} \cdot \boldsymbol{w} \in L^2(\Omega)$. Since Ω is a bounded domain, it follows that $\boldsymbol{v} \cdot \boldsymbol{w} \in L^1(\Omega)$. Then, there is a constant C such that

$$\int_\Omega (\boldsymbol{v} \cdot \boldsymbol{w} + C)\ d\boldsymbol{x} = 0$$

and consequently $\boldsymbol{v} \cdot \boldsymbol{w} + C \in Q$. If $\boldsymbol{u}$ satisfies the second equation of (6.2), i.e., $\boldsymbol{u}$ is weakly divergence-free, and if $\boldsymbol{u} \cdot \boldsymbol{n} = 0$ on Γ, then it follows with integration by parts that

$$\begin{aligned} 0 &= (\nabla \cdot \boldsymbol{u}, \boldsymbol{v} \cdot \boldsymbol{w} + C) = (\nabla \cdot \boldsymbol{u}, \boldsymbol{v} \cdot \boldsymbol{w}) + \int_\Gamma C\boldsymbol{u} \cdot \boldsymbol{n}\, ds - (\boldsymbol{u}, \nabla C) \\ &= (\nabla \cdot \boldsymbol{u}, \boldsymbol{v} \cdot \boldsymbol{w}). \end{aligned} \tag{6.23}$$

From (6.22), one obtains

$$(\boldsymbol{u}, \nabla (\boldsymbol{v} \cdot \boldsymbol{w})) = 0$$

and inserting this identity in (6.11) gives for $\boldsymbol{u}$ being weakly divergence-free with $\boldsymbol{u} \cdot \boldsymbol{n} = 0$ on Γ and $\boldsymbol{v}, \boldsymbol{w} \in H^1(\Omega)$

$$n_{\text{conv}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = -n_{\text{conv}}(\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v}). \tag{6.24}$$

With this relation, the skew-symmetric form of the convective term is defined by

$$n_{\text{skew}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = \frac{1}{2} \left(n_{\text{conv}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) - n_{\text{conv}}(\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v}) \right). \tag{6.25}$$

□

Remark 6.9 (Vanishing of the Convective Term) Analogously to the Oseen equations, the vanishing of the convective term in the case that the last two arguments are identical is important for the analysis and finite element error analysis. For a discussion of situations when this property is given, it is referred to Remark 5.6.

□

Lemma 6.10 (Vanishing of the Convective Term) *Let $\boldsymbol{u}, \boldsymbol{v} \in H^1(\Omega)$, then*

$$n_{\text{rot}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{v}) = n_{\text{skew}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{v}) = 0. \tag{6.26}$$

If additionally $\boldsymbol{u} \cdot \boldsymbol{n} = 0$ or $\boldsymbol{v} = \boldsymbol{0}$ on Γ, then

$$n_{\text{div}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{v}) = 0. \tag{6.27}$$

If $\boldsymbol{u}$ is weakly divergence-free and if $\boldsymbol{u} \cdot \boldsymbol{n} = 0$ on Γ, then

$$n_{\text{conv}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{v}) = 0. \tag{6.28}$$

Proof Property (6.26) for $n_{\text{skew}}(\cdot, \cdot, \cdot)$ follows directly from the definition. A direct calculation gives for the rotational form

$$\begin{aligned}
& n_{\text{rot}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{v}) \\
&= \left(\begin{pmatrix} v_3 \partial_z u_1 - v_3 \partial_x u_3 - v_2 \partial_x u_2 + v_2 \partial_y u_1 \\ v_1 \partial_x u_2 - v_1 \partial_y u_1 - v_3 \partial_y u_3 + v_3 \partial_z u_2 \\ v_2 \partial_y u_3 - v_2 \partial_z u_2 - v_1 \partial_z u_1 + v_1 \partial_x u_3 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \right) \\
&= \int_\Omega \Big[\partial_y u_1 (v_1 v_2 - v_1 v_2) + \partial_z u_1 (v_1 v_3 - v_1 v_3) + \partial_x u_2 (v_1 v_2 - v_1 v_2) \\
&\quad + \partial_z u_2 (v_2 v_3 - v_2 v_3) + \partial_x u_3 (v_1 v_3 - v_1 v_3) + \partial_y u_3 (v_2 v_3 - v_2 v_3) \Big] \, d\boldsymbol{x} \\
&= 0.
\end{aligned}$$

Considering the divergence form, then one gets with (6.11) and (6.22)

$$\begin{aligned} n_{\mathrm{div}}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{v}) &= n_{\mathrm{conv}}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{v}) + \frac{1}{2}\left((\nabla\cdot\boldsymbol{u})\,\boldsymbol{v},\boldsymbol{v}\right) \\ &= (\boldsymbol{u},\nabla(\boldsymbol{v}\cdot\boldsymbol{v})) - n_{\mathrm{conv}}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{v}) + \frac{1}{2}\left((\nabla\cdot\boldsymbol{u})\,\boldsymbol{v},\boldsymbol{v}\right) \\ &= -n_{\mathrm{conv}}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{v}) - (\nabla\cdot\boldsymbol{u},\boldsymbol{v}\cdot\boldsymbol{v}) + \frac{1}{2}\left((\nabla\cdot\boldsymbol{u})\,\boldsymbol{v},\boldsymbol{v}\right) \\ &= -n_{\mathrm{div}}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{v}), \end{aligned}$$

from what $n_{\mathrm{div}}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{v}) = 0$ follows.

For the convective form, (6.28) follows from (6.24). Note that the derivation of (6.24) used the assumptions stated in the formulation of the lemma. ■

Lemma 6.11 (Estimates of the Convective Term) *Let $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in H^1(\Omega)$, where $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, is a bounded domain with Lipschitz boundary, then there is a $C \in \mathbb{R}$ such that*

$$|n_{\mathrm{conv}}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})| \le C\,\|\boldsymbol{u}\|_{H^1(\Omega)}\,\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}\,\|\boldsymbol{w}\|_{H^1(\Omega)}\,. \tag{6.29}$$

For the skew-symmetric form, it holds

$$|n_{\mathrm{skew}}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})| \le C\,\|\boldsymbol{u}\|_{H^1(\Omega)}\,\|\boldsymbol{v}\|_{H^1(\Omega)}\,\|\boldsymbol{w}\|_{H^1(\Omega)} \tag{6.30}$$

and for the divergence form of the convective term

$$|n_{\mathrm{div}}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})| \le C\,\|\boldsymbol{u}\|_{H^1(\Omega)}\,\|\boldsymbol{v}\|_{H^1(\Omega)}\,\|\boldsymbol{w}\|_{H^1(\Omega)}\,. \tag{6.31}$$

If $\boldsymbol{w} \in V_{\mathrm{div}}$, then it is for the rotational form

$$|n_{\mathrm{rot}}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})| \le C\,\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}\,\|\boldsymbol{v}\|_{H^1(\Omega)}\,\|\boldsymbol{w}\|_{H^1(\Omega)}\,. \tag{6.32}$$

Proof The estimate starts with the application of the generalized Hölder inequality

$$|n_{\mathrm{conv}}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})| = \left|\int_\Omega (\boldsymbol{u}\cdot\nabla)\,\boldsymbol{v}\cdot\boldsymbol{w}\,d\boldsymbol{x}\right| \le \|\boldsymbol{u}\|_{L^p(\Omega)}\,\|\nabla\boldsymbol{v}\|_{L^q(\Omega)}\,\|\boldsymbol{w}\|_{L^r(\Omega)}\,, \tag{6.33}$$

with

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1,\ \ 1 \le p,q,r \le \infty.$$

Since $\boldsymbol{v} \in H^1(\Omega)$, one can take at most $q = 2$. The other two terms are of the same form such that they can treated similarly, i.e., one can take $p = r = 4$.

Applying the Sobolev imbedding $H^1(\Omega) \to L^4(\Omega)$, see (A.21) and (A.22), gives immediately (6.29).

The statement for the skew-symmetric term follows by applying the triangle inequality

$$|n_{\text{skew}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})| \le \frac{1}{2} \left(|n_{\text{conv}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})| + |n_{\text{conv}}(\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v})| \right)$$

and then estimate for the convective term.

For the divergence form of the convective term, one gets for the second part of this term, using the generalized Hölder's inequality, the Sobolev imbedding $H^1(\Omega) \to L^4(\Omega)$, and (3.40)

$$\begin{aligned} \frac{1}{2} \left((\nabla \cdot \boldsymbol{u})\, \boldsymbol{v}, \boldsymbol{w} \right) &\le \frac{1}{2} \|\nabla \cdot \boldsymbol{u}\|_{L^2(\Omega)} \|\boldsymbol{v}\|_{L^4(\Omega)} \|\boldsymbol{w}\|_{L^4(\Omega)} \\ &\le C \|\nabla \cdot \boldsymbol{u}\|_{L^2(\Omega)} \|\boldsymbol{v}\|_{H^1(\Omega)} \|\boldsymbol{w}\|_{H^1(\Omega)} \\ &\le C \|\nabla \boldsymbol{u}\|_{L^2(\Omega)} \|\boldsymbol{v}\|_{H^1(\Omega)} \|\boldsymbol{w}\|_{H^1(\Omega)} . \end{aligned}$$

Since the first part of the divergence form is just the convective form, estimate (6.31) follows by combining the estimate for the second part and (6.29).

If $\boldsymbol{w} \in V_{\text{div}}$, then it follows from (6.14) that

$$n_{\text{rot}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = n_{\text{conv}}(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{w}) - n_{\text{conv}}(\boldsymbol{w}, \boldsymbol{u}, \boldsymbol{v}). \tag{6.34}$$

Now, (6.32) is proved by applying the triangle inequality and (6.29). ∎

Remark 6.12 (On the Convective Term)

- If $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$, then the application of the Poincaré–Friedrichs inequality (A.12) to (6.29)–(6.31) gives

$$|n_{\text{conv}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})| \le C(\Omega) \|\boldsymbol{u}\|_V \|\boldsymbol{v}\|_V \|\boldsymbol{w}\|_V , \tag{6.35}$$

$$|n_{\text{skew}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})| \le C(\Omega) \|\boldsymbol{u}\|_V \|\boldsymbol{v}\|_V \|\boldsymbol{w}\|_V , \tag{6.36}$$

$$|n_{\text{div}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})| \le C(\Omega) \|\boldsymbol{u}\|_V \|\boldsymbol{v}\|_V \|\boldsymbol{w}\|_V . \tag{6.37}$$

- In the proof of Lemma 6.11, other choices of the parameters in Hölder's inequality are possible, which may lead to different estimates of the trilinear term. Further estimates can be derived for different regularity assumptions on the functions, e.g., see Layton and Tobiska (1998). Several estimates of this kind, which will be used in this monograph, are derived below.

□

Lemma 6.13 (Estimates of the Convective Term) *Let* $\boldsymbol{u} \in L^2(\Omega)$, $\boldsymbol{v} \in W^{1,\infty}(\Omega)$, *and* $\boldsymbol{w} \in H^1(\Omega)$, *then*

$$|n_{\text{conv}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})| \le \|\boldsymbol{u}\|_{L^2(\Omega)} \|\nabla \boldsymbol{v}\|_{L^\infty(\Omega)} \|\boldsymbol{w}\|_{L^2(\Omega)} \tag{6.38}$$

and

$$|n_{\text{skew}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})| \le \frac{1}{2}\Bigg(\|\boldsymbol{u}\|_{L^2(\Omega)} \|\nabla \boldsymbol{v}\|_{L^\infty(\Omega)} \|\boldsymbol{w}\|_{L^2(\Omega)} + \|\boldsymbol{u}\|_{L^2(\Omega)} \|\nabla \boldsymbol{w}\|_{L^2(\Omega)} \|\boldsymbol{v}\|_{L^\infty(\Omega)} \Bigg). \tag{6.39}$$

In particular, it is

$$|n_{\text{skew}}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})| \le C \|\boldsymbol{u}\|_{L^2(\Omega)} \|\boldsymbol{v}\|_{W^{1,\infty}(\Omega)} \|\boldsymbol{w}\|_{H^1(\Omega)} . \tag{6.40}$$

Proof Estimates (6.38) and (6.39) follow directly from the generalized Hölder inequality (6.33) by choosing $p = r = 2, q = \infty$ and $p = q = 2, r = \infty$, respectively. Using the definitions of the norms in estimate (6.40) shows that this estimate is a direct consequence of estimate (6.39). ■

Lemma 6.14 (Estimates of the Convective Term) *Let* $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$, *then it holds*

$$n(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) \le C \|\boldsymbol{u}\|_{L^2(\Omega)}^{1-s} \|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^{s} \|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \|\nabla \boldsymbol{w}\|_{L^2(\Omega)} \tag{6.41}$$

with arbitrary $s \in (0, 1]$ *if* $d = 2$ *and* $s \in [1/2, 1]$ *if* $d = 3$. *In* (6.41), *the trilinear forms* $n_{\text{conv}}(\cdot,\cdot,\cdot)$ *and* $n_{\text{skew}}(\cdot,\cdot,\cdot)$ *can be used.*

Proof Consider the convective form of the nonlinear term and $d = 2$. Then, one obtains with the generalized Hölder inequality (6.33), the Sobolev imbedding (A.21) (with q arbitrarily large, $p = 2 + \varepsilon$, $\varepsilon > 0$), Poincaré's inequality (A.12), and the Sobolev imbedding (A.15) (with $p = 2, q = 2 + \varepsilon, m = \varepsilon/(2 + \varepsilon) = s$ for arbitrary $\varepsilon > 0$)

$$\begin{aligned}((\boldsymbol{u} \cdot \nabla) \boldsymbol{v}, \boldsymbol{w}) &\le \|\boldsymbol{u}\|_{L^p(\Omega)} \|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \|\boldsymbol{w}\|_{L^q(\Omega)} \quad \text{with } p^{-1} + q^{-1} = 1/2 \\ &\le C \|\boldsymbol{u}\|_{L^{2+\varepsilon}(\Omega)} \|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \|\nabla \boldsymbol{w}\|_{L^2(\Omega)} \\ &\le C \|\boldsymbol{u}\|_{H^s(\Omega)} \|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \|\nabla \boldsymbol{w}\|_{L^2(\Omega)} .\end{aligned}$$

Finally, the interpolation estimate (A.13) is applied to the first factor on the right-hand side, followed by Poincaré's inequality, such that

$$((\boldsymbol{u} \cdot \nabla) \boldsymbol{v}, \boldsymbol{w}) \le C \|\boldsymbol{u}\|_{L^2(\Omega)}^{1-s} \|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^{s} \|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \|\nabla \boldsymbol{w}\|_{L^2(\Omega)}$$

for arbitrary small positive s.

For $d = 3$, the way to prove the estimate is similar, it will be presented in detail only for $s = 1/2$. After having applied Hölder's inequality, the Sobolev imbedding (A.22), the Sobolev imbedding (A.15) with $p = 2, q = 3, m = 1/2$, the interpolation estimate (A.13), and Poincaré's inequality, one obtains

$$\begin{aligned}((\boldsymbol{u}\cdot\nabla)\,\boldsymbol{v},\boldsymbol{w}) &\le \|\boldsymbol{u}\|_{L^3(\Omega)}\,\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}\,\|\boldsymbol{w}\|_{L^6(\Omega)}\\ &\le C\,\|\boldsymbol{u}\|_{L^3(\Omega)}\,\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}\,\|\nabla\boldsymbol{w}\|_{L^2(\Omega)}\\ &\le C\,\|\boldsymbol{u}\|_{H^{1/2}(\Omega)}\,\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}\,\|\nabla\boldsymbol{w}\|_{L^2(\Omega)}\\ &\le C\,\|\boldsymbol{u}\|_{L^2(\Omega)}^{1/2}\,\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^{1/2}\,\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}\,\|\nabla\boldsymbol{w}\|_{L^2(\Omega)}\,.\end{aligned}$$

The estimate for the skew-symmetric term follows from applying the estimate of the convective term to both parts of the skew-symmetric term. ■

Lemma 6.15 (Estimates of the Convective Term) *Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$, then it holds*

$$n_{\text{conv}}\,(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) \le C\,\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}\,\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}\,\|\boldsymbol{w}\|_{L^2(\Omega)}^{1-s}\,\|\nabla\boldsymbol{w}\|_{L^2(\Omega)}^{s} \tag{6.42}$$

with arbitrary $s \in (0,1]$ if $d = 2$ and $s \in [1/2,1]$ if $d = 3$.

If in addition $\boldsymbol{w} \in V_{\text{div}}$, then it is

$$|n_{\text{rot}}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})| \le C\,\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}\,\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}\,\|\boldsymbol{w}\|_{L^2(\Omega)}^{1-s}\,\|\nabla\boldsymbol{w}\|_{L^2(\Omega)}^{s}\,, \tag{6.43}$$

$$|n_{\text{rot}}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})| \le C\,\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}\,\|\boldsymbol{v}\|_{L^2(\Omega)}^{1-s}\,\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}^{s}\,\|\nabla\boldsymbol{w}\|_{L^2(\Omega)}\,. \tag{6.44}$$

Proof Estimate (6.42) is derived with the same tools as estimate (6.41), just changing the roles of p and q in the proof.

The statements for the rotational form follow from (6.34), the triangle inequality, and estimates (6.41) and (6.42). ■

6.1.3 Existence, Uniqueness, and Stability of a Solution

Remark 6.16 (Contents) This section presents results on the existence, uniqueness, and stability of a weak solution of the Navier–Stokes equations (6.2). It turns out that a solution always exists but it is unique only in the case of sufficiently small external forces and sufficiently large viscosity, see (6.48) below for the concrete requirement.

From the point of view of numerical simulations, the uniqueness case is the only interesting one. In the non-uniqueness case, small perturbations of the data will lead to time-dependent solutions. From the practical point of view, one should consider and discretize in such a case the time-dependent Navier–Stokes equations. □

Theorem 6.17 (Existence of a Solution) *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be a bounded domain with Lipschitz boundary $\partial\Omega$ and let $\boldsymbol{f} \in H^{-1}(\Omega)$. Then there exists at least one solution of* (6.2).

Proof For the proof it is referred to Girault and Raviart (1986, Chap. IV, Theorems 2.3 and 2.4), Galdi (2011, Chap. IX.3), or Temam (1984, Chap. II, Theorem 1.2). Essential ideas of the proof are as follows:

- The equivalent problem (6.4) in the divergence-free subspace V_{div} is considered, such that only the velocity appears.
- The problem is considered in finite-dimensional spaces (Galerkin method).
- A fixed point equation is constructed and the existence of a solution of the finite-dimensional problems is proved by the fixed point theorem of Brouwer, Theorem A.69.
- It is shown that for the dimension of the spaces going to infinity, a subsequence of the solutions tends to a solution of problem (6.4).
- The existence of the pressure is recovered with the help of the inf-sup condition (3.14).

■

Remark 6.18 (Norms of the Trilinear Form of the Convective Term) The norm of the convective term is denoted by

$$N = \sup_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in V\setminus\{\boldsymbol{0}\}} \frac{((\boldsymbol{u}\cdot\nabla)\boldsymbol{v},\boldsymbol{w})}{\|\boldsymbol{u}\|_V\,\|\boldsymbol{v}\|_V\,\|\boldsymbol{w}\|_V}.$$

Note that N is the smallest constant in estimate (6.35) since

$$N \ge \frac{((\boldsymbol{u}\cdot\nabla)\boldsymbol{v},\boldsymbol{w})}{\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}\,\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}\,\|\nabla\boldsymbol{w}\|_{L^2(\Omega)}} \quad \forall\ \boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in V,$$

which is the same inequality as (6.35). The existence of a smaller constant in (6.35) than N contradicts the definition of N. Likewise, define

$$N_{\text{div}} = \sup_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in V_{\text{div}}\setminus\{\boldsymbol{0}\}} \frac{((\boldsymbol{u}\cdot\nabla)\boldsymbol{v},\boldsymbol{w})}{\|\boldsymbol{u}\|_V\,\|\boldsymbol{v}\|_V\,\|\boldsymbol{w}\|_V}. \tag{6.45}$$

Since $V_{\text{div}} \subset V$ and the supremum in a subset cannot be larger than the supremum in the whole set, if follows that $0 < N_{\text{div}} \le N < \infty$. The boundedness of N is a consequence of estimate (6.35). □

Remark 6.19 (Definition of an Operator with the Help of an Oseen Problem) Let $\boldsymbol{b} \in V_{\text{div}}$ and consider the Oseen problem: Find $\boldsymbol{u} \in V_{\text{div}}$ such that

$$(\nu\nabla\boldsymbol{u},\nabla\boldsymbol{v}) + ((\boldsymbol{b}\cdot\nabla)\,\boldsymbol{u},\boldsymbol{v}) = \langle\boldsymbol{f},\boldsymbol{v}\rangle_{V',V} \quad \forall\ \boldsymbol{v}\in V_{\text{div}}. \tag{6.46}$$

This problem has a unique solution, see Theorem 5.7, and this solution satisfies the stability estimate

$$\|\nabla \boldsymbol{u}\|_{L^2(\Omega)} \le \frac{1}{\nu} \|\boldsymbol{f}\|_{H^{-1}(\Omega)}, \tag{6.47}$$

compare (5.9).

With the Oseen problem (6.46), an operator is defined that maps the given convection field $\boldsymbol{b}$ to the solution $\boldsymbol{u}$

$$N_{\text{conv}} \ : \ V_{\text{div}} \to V_{\text{div}} \quad \boldsymbol{b} \mapsto \boldsymbol{u}.$$

It is obvious that each fixed point $\boldsymbol{u}_*$ of N_{conv} is a velocity solution of the Navier–Stokes equations (6.4), since then it follows from (6.46) that

$$(\nu \nabla \boldsymbol{u}_*, \nabla \boldsymbol{v}) + ((\boldsymbol{u}_* \cdot \nabla) \boldsymbol{u}_*, \boldsymbol{v}) = \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{V',V} \quad \forall\ \boldsymbol{v} \in V_{\text{div}}.$$

□

Theorem 6.20 (Existence and Uniqueness of a Solution for Small Data) *Let the assumptions of Theorem 6.17 be satisfied and let in addition*

$$\frac{N_{\text{div}} \|\boldsymbol{f}\|_{H^{-1}(\Omega)}}{\nu^2} < 1, \tag{6.48}$$

then problem (6.4) has a unique solution $\boldsymbol{u} \in V_{\text{div}}$ *and problem (6.2) has a unique solution* $(\boldsymbol{u}, p) \in V \times Q$.

Proof It will be shown that N_{conv} defines a contraction on V_{div}. First of all, it can be observed that N_{conv} is bounded independently of $\boldsymbol{b}$, since one obtains with (6.47)

$$\begin{aligned} \|N_{\text{conv}}\| &= \sup_{\boldsymbol{b} \in V_{\text{div}}, \|\boldsymbol{b}\|_V = 1} \|N_{\text{conv}} \boldsymbol{b}\|_V = \sup_{\boldsymbol{b} \in V_{\text{div}}, \|\boldsymbol{b}\|_V = 1} \|\boldsymbol{u}\|_V \\ &\le \sup_{\boldsymbol{b} \in V_{\text{div}}, \|\boldsymbol{b}\|_V = 1} \frac{1}{\nu} \|\boldsymbol{f}\|_{H^{-1}(\Omega)} = \frac{1}{\nu} \|\boldsymbol{f}\|_{H^{-1}(\Omega)}. \end{aligned}$$

Now, one chooses $\boldsymbol{b}_1, \boldsymbol{b}_2 \in V_{\text{div}}$ arbitrarily and denotes $\boldsymbol{u}_1 = N_{\text{conv}} \boldsymbol{b}_1, \boldsymbol{u}_2 = N_{\text{conv}} \boldsymbol{b}_2$. Both functions $\boldsymbol{u}_1, \boldsymbol{u}_2$ solve the Oseen equation (6.46) with the same right-hand side. Subtracting these equations, one gets

$$\begin{aligned} 0 &= (\nu \nabla \boldsymbol{u}_1, \nabla \boldsymbol{v}) + ((\boldsymbol{b}_1 \cdot \nabla) \boldsymbol{u}_1, \boldsymbol{v}) - (\nu \nabla \boldsymbol{u}_2, \nabla \boldsymbol{v}) - ((\boldsymbol{b}_2 \cdot \nabla) \boldsymbol{u}_2, \boldsymbol{v}) \\ &= \nu (\nabla (\boldsymbol{u}_1 - \boldsymbol{u}_2), \nabla \boldsymbol{v}) + (((\boldsymbol{b}_1 - \boldsymbol{b}_2) \cdot \nabla) \boldsymbol{u}_1, \boldsymbol{v}) + ((\boldsymbol{b}_2 \cdot \nabla) (\boldsymbol{u}_1 - \boldsymbol{u}_2), \boldsymbol{v}) \\ &\quad \forall\ \boldsymbol{v} \in V_{\text{div}}. \end{aligned}$$

Setting $\boldsymbol{v} = \boldsymbol{u}_1 - \boldsymbol{u}_2 \in V_{\text{div}}$, the last term on the right-hand side vanishes because of (6.28) and one obtains with (6.45), (6.47), and (6.48)

$$\begin{aligned}\|\nabla (\boldsymbol{u}_1 - \boldsymbol{u}_2)\|_{L^2(\Omega)}^2 = \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_V^2 &= -\frac{1}{\nu}\left(\left(\left(\boldsymbol{b}_1 - \boldsymbol{b}_2\right)\cdot \nabla\right)\boldsymbol{u}_1, \boldsymbol{u}_1 - \boldsymbol{u}_2\right) \\ &\le \frac{N_{\text{div}}}{\nu}\|\boldsymbol{b}_1 - \boldsymbol{b}_2\|_V \|\boldsymbol{u}_1\|_V \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_V \\ &\le \frac{N_{\text{div}}\|\boldsymbol{f}\|_{H^{-1}(\Omega)}}{\nu^2}\|\boldsymbol{b}_1 - \boldsymbol{b}_2\|_V \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_V \\ &< \|\boldsymbol{b}_1 - \boldsymbol{b}_2\|_V \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_V .\end{aligned}$$

It follows that

$$\|N_{\text{conv}}\boldsymbol{b}_1 - N_{\text{conv}}\boldsymbol{b}_2\|_V = \|\boldsymbol{u}_1 - \boldsymbol{u}_2\|_V < \|\boldsymbol{b}_1 - \boldsymbol{b}_2\|_V \quad \forall\ \boldsymbol{b}_1, \boldsymbol{b}_2 \in V_{\text{div}},$$

which is the contraction property for N_{conv}. The existence and uniqueness of a solution of problem (6.4) follows now with the fixed point theorem of Banach, Theorem A.68. The uniqueness of the solution of problem (6.2) is a consequence of the fact that V and Q satisfy the inf-sup condition (3.14), see Theorem 3.46. ■

Lemma 6.21 (Stability of the Solution) *Let* $(\boldsymbol{u}, p) \in V \times Q$ *be any solution of* (6.2)*, then*

$$\|\nabla \boldsymbol{u}\|_{L^2(\Omega)} \le \frac{1}{\nu}\|\boldsymbol{f}\|_{H^{-1}(\Omega)}, \tag{6.49}$$

$$\|p\|_{L^2(\Omega)} \le \frac{1}{\beta_{\text{is}}}\left(2\|\boldsymbol{f}\|_{H^{-1}(\Omega)} + \frac{C}{\nu^2}\|\boldsymbol{f}\|_{H^{-1}(\Omega)}^2\right). \tag{6.50}$$

Proof The proof starts in the usual way by choosing as test function $(\boldsymbol{v}, q) = (\boldsymbol{u}, p)$ in (6.3)

$$(\nu\nabla\boldsymbol{u}, \nabla\boldsymbol{u}) + n(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{u}) = \langle \boldsymbol{f}, \boldsymbol{u}\rangle_{V',V},$$

where $n(\cdot,\cdot,\cdot)$ is any of the convective terms introduced in Remark 6.8. With (6.26)–(6.28), it follows that

$$\nu\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^2 = \langle \boldsymbol{f}, \boldsymbol{u}\rangle_{V',V}.$$

The application of the inequality for the dual pairing gives

$$\nu\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^2 \le \|\boldsymbol{f}\|_{H^{-1}(\Omega)}\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}.$$

This inequality is equivalent to (6.49).

Starting with the inf-sup condition, one obtains for the pressure with inserting (6.3), the estimate for the dual pairing, the Cauchy–Schwarz inequality (A.10), and (6.35)

$$\begin{aligned}\|p\|_{L^2(\Omega)} &\leq \frac{1}{\beta_{\text{is}}} \sup_{\boldsymbol{v}\in V\setminus\{\boldsymbol{0}\}} \frac{-(\nabla\cdot\boldsymbol{v},p)}{\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}} \\ &= \frac{1}{\beta_{\text{is}}} \sup_{\boldsymbol{v}\in V\setminus\{\boldsymbol{0}\}} \frac{(\boldsymbol{f},\boldsymbol{v})-(\nu\nabla\boldsymbol{u},\nabla\boldsymbol{v})-n(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v})}{\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}} \\ &\leq \frac{1}{\beta_{\text{is}}}\left(\|\boldsymbol{f}\|_{H^{-1}(\Omega)}+\nu\,\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}+C\,\|\nabla\boldsymbol{u}\|^2_{L^2(\Omega)}\right).\end{aligned}$$

The statement follows now by inserting the stability estimate (6.49) for the velocity. ■

Remark 6.22 (Directional Do-Nothing Condition) The Navier–Stokes equations equipped with the directional do-nothing condition (2.41) on some part of the domain are analyzed in Braack and Mucha (2014). In particular, the existence of a weak solution is proved. □

6.2 The Galerkin Finite Element Method

Remark 6.23 (Contents) This section discusses finite element error estimates. The essential approaches were already presented for the Galerkin discretizations of the Stokes and the Oseen problem. For the steady-state Navier–Stokes equations, in addition an estimate of the nonlinear (trilinear) term is necessary. □

Remark 6.24 (The Galerkin Finite Element Formulation of the Steady-State Navier–Stokes Equations) Let $V^h \subset V$ and $Q^h \subset Q$ be inf-sup stable finite element spaces, i.e., (3.51) is fulfilled. Then, the Galerkin finite element discretization of the steady-state Navier–Stokes equations reads as follows: Find $(\boldsymbol{u}^h, p^h) \in V^h \times Q^h$ such that

$$\begin{aligned}\nu\left(\nabla\boldsymbol{u}^h,\nabla\boldsymbol{v}^h\right)+n\left(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{v}^h\right)-\left(\nabla\cdot\boldsymbol{v}^h,p^h\right) &= \langle\boldsymbol{f},\boldsymbol{v}^h\rangle_{V',V} \quad \forall\ \boldsymbol{v}^h\in V^h, \\ -\left(\nabla\cdot\boldsymbol{u}^h,q^h\right) &= 0 \qquad\qquad\quad \forall\ q^h\in Q^h,\end{aligned} \tag{6.51}$$

where $n(\cdot,\cdot,\cdot)$ is any of the convective terms introduced in Remark 6.8. An equivalent formulation is as follows: Find $(\boldsymbol{u}^h, p^h) \in V^h \times Q^h$ such that

$$\nu\left(\nabla\boldsymbol{u}^h,\nabla\boldsymbol{v}^h\right)+n\left(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{v}^h\right)-\left(\nabla\cdot\boldsymbol{v}^h,p^h\right)+\left(\nabla\cdot\boldsymbol{u}^h,q^h\right) = \langle\boldsymbol{f},\boldsymbol{v}^h\rangle_{V',V} \tag{6.52}$$

for all $(\boldsymbol{v}^h, q^h) \in V^h \times Q^h$.

By the satisfaction of the discrete inf-sup condition, it is known that the space V^h_{div} is not empty, see Remark 3.16.

Existence and uniqueness of a solution are proved in the similar way as for the continuous equation. In particular, a unique solution can be expected only for small external forces (right-hand side) and a large viscosity. A discrete analog to condition (6.48) will be assumed throughout this section. □

Remark 6.25 (On the Convective Term for the Finite Element Error Analysis) A main tool of the analysis of the Oseen equations is property (5.4) that states that the convective term vanishes if the convection field is (weakly) divergence-free and the ansatz and test functions are identical. A similar property for the different types of the nonlinear convective term of the continuous Navier–Stokes equations was proved in Lemma 6.10.

For $n_{\mathrm{conv}}(\cdot,\cdot,\cdot)$, the proof of this property relies on the assumption that the convection field is weakly divergence-free. Since generally $V_{\mathrm{div}}^h \not\subset V_{\mathrm{div}}$, see Remark 3.56, finite element velocity fields will be usually not weakly divergence-free and the convective term $n_{\mathrm{conv}}\left(\boldsymbol{u}^h, \boldsymbol{v}^h, \boldsymbol{v}^h\right)$ will not vanish for $\boldsymbol{u}^h \in V_{\mathrm{div}}^h$ and $\boldsymbol{v}^h \in V^h$.

In contrast, it is obvious to observe that

$$n_{\mathrm{skew}}\left(\boldsymbol{u}^h, \boldsymbol{v}^h, \boldsymbol{v}^h\right) = \frac{1}{2}\left(n_{\mathrm{conv}}\left(\boldsymbol{u}^h, \boldsymbol{v}^h, \boldsymbol{v}^h\right) - n_{\mathrm{conv}}\left(\boldsymbol{u}^h, \boldsymbol{v}^h, \boldsymbol{v}^h\right)\right) = 0 \tag{6.53}$$

for all $\boldsymbol{u}^h, \boldsymbol{v}^h \in V^h$. Analogously to the proof of Lemma 6.10, one finds by direct computations that also

$$n_{\mathrm{rot}}\left(\boldsymbol{u}^h, \boldsymbol{v}^h, \boldsymbol{v}^h\right) = 0 \quad \forall\ \boldsymbol{u}^h, \boldsymbol{v}^h \in V^h \tag{6.54}$$

and

$$n_{\mathrm{div}}\left(\boldsymbol{u}^h, \boldsymbol{v}^h, \boldsymbol{v}^h\right) = 0 \quad \forall\ \boldsymbol{u}^h, \boldsymbol{v}^h \in V^h. \tag{6.55}$$

Main tools in the finite element error analysis of the convective term are properties of the form (6.53)–(6.55), and estimates of the convective term from above, see (6.36) and (6.37).

Usually, one finds in the literature that the analysis is carried out for the skew-symmetric form of the convective term. The presentation below will also use this form. □

Theorem 6.26 (Existence and Uniqueness of the Galerkin Finite Element Solution for Small Data) *Let the assumptions of Theorem 6.17 be satisfied and let in addition*

$$\frac{N_{\mathrm{div}}^h \|\boldsymbol{f}\|_{H^{-1}(\Omega)}}{\nu^2} < 1, \tag{6.56}$$

where

$$N_{\mathrm{div}}^h = \sup_{\boldsymbol{u}^h,\boldsymbol{v}^h,\boldsymbol{w}^h\in V_{\mathrm{div}}^h\setminus\{\boldsymbol{0}\}} \frac{\frac{1}{2}\left(((\boldsymbol{u}^h\cdot\nabla)\boldsymbol{v}^h,\boldsymbol{w}^h)-((\boldsymbol{u}^h\cdot\nabla)\boldsymbol{w}^h,\boldsymbol{v}^h)\right)}{\|\boldsymbol{u}^h\|_V\,\|\boldsymbol{v}^h\|_V\,\|\boldsymbol{w}^h\|_V}. \tag{6.57}$$

If a pair of finite element spaces $V^h \times Q^h$ is used that satisfies the discrete inf-sup condition (3.51)*, then problem* (6.51) *with $n\left(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{v}^h\right) = n_{\mathrm{skew}}\left(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{v}^h\right)$ has a unique solution $(\boldsymbol{u}^h,p^h)\in V^h\times Q^h$.*

Proof The proof follows the lines for proving the uniqueness of the solution for the continuous problem, see Theorem 6.20. One considers now an Oseen problem in V_{div}^h: Given $\boldsymbol{b}^h \in V_{\mathrm{div}}^h$, find $\boldsymbol{u}^h \in V_{\mathrm{div}}^h$ such that for given $\boldsymbol{f}\in V'$

$$\left(\nu\nabla\boldsymbol{u}^h,\nabla\boldsymbol{v}^h\right)+\frac{1}{2}\left(\left(\left(\boldsymbol{b}^h\cdot\nabla\right)\boldsymbol{u}^h,\boldsymbol{v}^h\right)-\left(\left(\boldsymbol{b}^h\cdot\nabla\right)\boldsymbol{v}^h,\boldsymbol{u}^h\right)\right)=\langle\boldsymbol{f},\boldsymbol{v}\rangle_{V',V}\quad\forall\ \boldsymbol{v}\in V_{\mathrm{div}}^h.$$

Analogously to Corollary 5.12, the existence of a unique solution of this problem is proved and a stability estimate of the form (5.14) is derived as in Lemma 5.13. Note that for both statements it is not of importance whether the convection field is weakly divergence-free or only discretely divergence-free, but only the vanishing of the convective term is used in the case that the last two arguments are identical. Then, one defines a linear operator $N_{\mathrm{conv}}^h : V_{\mathrm{div}}^h \to V_{\mathrm{div}}^h$ which maps $\boldsymbol{b}^h \to \boldsymbol{u}^h$. Now, one shows, analogously to the proof of Theorem 6.20, that N_{conv}^h is bounded and that it is a contraction. ■

Lemma 6.27 (Stability of the Finite Element Solution) *Let $V^h \times Q^h$ be a pair of inf-sup stable finite element spaces. Then, the finite element solution of the steady-state Navier–Stokes equations with skew-symmetric form of the convective term is stable*

$$\left\|\nabla\boldsymbol{u}^h\right\|_{L^2(\Omega)} \le \frac{1}{\nu}\,\|\boldsymbol{f}\|_{H^{-1}(\Omega)}\,, \tag{6.58}$$

$$\left\|p^h\right\|_{L^2(\Omega)} \le \frac{1}{\beta_{\mathrm{is}}^h}\left(2\,\|\boldsymbol{f}\|_{H^{-1}(\Omega)}+\frac{C}{\nu^2}\,\|\boldsymbol{f}\|_{H^{-1}(\Omega)}^2\right). \tag{6.59}$$

Proof The proof is performed analogously to the proof of Lemma 6.21. ■

Theorem 6.28 (Finite Element Error Estimate for the $L^2(\Omega)$ Norm of the Gradient of the Velocity) *Let $\Omega\subset\mathbb{R}^d$, $d\in\{2,3\}$, be a bounded domain with polyhedral and Lipschitz continuous boundary, let* (6.48) *be fulfilled, and let instead of* (6.56) *the stronger condition*

$$\frac{N_{\mathrm{div}}^h\,\|\boldsymbol{f}\|_{H^{-1}(\Omega)}}{\nu^2} \le \frac{1}{4}, \tag{6.60}$$

be satisfied. Let $(\boldsymbol{u}, p) \in V \times Q$ *be the unique solution of the Navier–Stokes equations* (6.2)*. Assume that this problem is discretized with inf-sup stable finite element spaces* $V^h \times Q^h$ *using the skew-symmetric form of the convective term and denote by* $\boldsymbol{u}^h \in V^h_{\mathrm{div}}$ *the velocity solution. Then, the following error estimate holds*

$$\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)} \le C\left(\left(1+\frac{1}{\nu^2}\|\boldsymbol{f}\|_{H^{-1}(\Omega)}\right)\inf_{\boldsymbol{v}^h\in V^h_{\mathrm{div}}}\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{v}^h\right)\right\|_{L^2(\Omega)} + \frac{1}{\nu}\inf_{q^h\in Q^h}\left\|p-q^h\right\|_{L^2(\Omega)}\right). \tag{6.61}$$

The constant C does not depend on the mesh size.

Proof The principle of the proof is the same as for the Stokes equations, see Theorem 4.21. Since the space V^h_{div} is not empty, one can use test functions $\boldsymbol{v}^h \in V^h_{\mathrm{div}}$ in (6.3) and (6.52) and subtract these equations to obtain the following error equation

$$\nu\left(\nabla(\boldsymbol{u}-\boldsymbol{u}^h),\nabla\boldsymbol{v}^h\right) + n_{\mathrm{skew}}\left(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}^h\right) - n_{\mathrm{skew}}\left(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{v}^h\right) - \left(\nabla\cdot\boldsymbol{v}^h, p-q^h\right) = 0$$

for all $\boldsymbol{v}^h \in V^h_{\mathrm{div}}$ and all $q^h \in Q^h$. In this step, also $\left(\nabla\cdot\boldsymbol{v}^h, q^h\right)$ for all $q^h \in Q^h$ was applied. Next, the error is decomposed in an approximation error and a discrete remainder

$$\boldsymbol{u}-\boldsymbol{u}^h = \left(\boldsymbol{u}-I^h\boldsymbol{u}\right) - \left(\boldsymbol{u}^h - I^h\boldsymbol{u}\right) = \boldsymbol{\eta}-\boldsymbol{\phi}^h, \quad I^h\boldsymbol{u}\in V^h_{\mathrm{div}}. \tag{6.62}$$

Inserting this decomposition in the error equation and setting $\boldsymbol{v}^h = \boldsymbol{\phi}^h$ leads to

$$\begin{aligned}\nu\left\|\nabla\boldsymbol{\phi}^h\right\|^2_{L^2(\Omega)} &= \nu\left(\nabla\boldsymbol{\eta},\nabla\boldsymbol{\phi}^h\right) - \left(\nabla\cdot\boldsymbol{\phi}^h, p-q^h\right) \\ &\quad + n_{\mathrm{skew}}\left(\boldsymbol{u},\boldsymbol{u},\boldsymbol{\phi}^h\right) - n_{\mathrm{skew}}\left(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{\phi}^h\right) \quad \forall\, q^h\in Q^h.\end{aligned} \tag{6.63}$$

The first two terms are estimated in a similar way as for the Oseen equations, compare the proof of Theorem 5.14,

$$\nu\left(\nabla\boldsymbol{\eta},\nabla\boldsymbol{\phi}^h\right) \le 2\nu\left\|\nabla\boldsymbol{\eta}\right\|^2_{L^2(\Omega)} + \frac{\nu}{8}\left\|\nabla\boldsymbol{\phi}^h\right\|^2_{L^2(\Omega)},$$

and

$$\left(\nabla\cdot\boldsymbol{\phi}^h, p-q^h\right) \le \frac{2}{\nu}\left\|p-q^h\right\|^2_{L^2(\Omega)} + \frac{\nu}{8}\left\|\nabla\boldsymbol{\phi}^h\right\|^2_{L^2(\Omega)}.$$

The new aspect for the Navier–Stokes equations is the estimate of the trilinear terms. Such terms are written in the form

$$aa - bb = aa - ab + ab - bb = a(a-b) + (a-b)b.$$

The differences are used to introduce approximation errors in the estimate. Applying this approach and using (6.62) yields

$$\begin{aligned}
& n_{\text{skew}}\left(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\boldsymbol{u}^h, \boldsymbol{u}^h, \boldsymbol{\phi}^h\right) \\
&= n_{\text{skew}}\left(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\boldsymbol{u}, \boldsymbol{u}^h, \boldsymbol{\phi}^h\right) + n_{\text{skew}}\left(\boldsymbol{u}, \boldsymbol{u}^h, \boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\boldsymbol{u}^h, \boldsymbol{u}^h, \boldsymbol{\phi}^h\right) \\
&= n_{\text{skew}}\left(\boldsymbol{u}, \boldsymbol{u} - \boldsymbol{u}^h, \boldsymbol{\phi}^h\right) + n_{\text{skew}}\left(\boldsymbol{u} - \boldsymbol{u}^h, \boldsymbol{u}^h, \boldsymbol{\phi}^h\right) \qquad (6.64) \\
&= n_{\text{skew}}\left(\boldsymbol{u}, \boldsymbol{\eta}, \boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\boldsymbol{u}, \boldsymbol{\phi}^h, \boldsymbol{\phi}^h\right) + n_{\text{skew}}\left(\boldsymbol{\eta}, \boldsymbol{u}^h, \boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\boldsymbol{\phi}^h, \boldsymbol{u}^h, \boldsymbol{\phi}^h\right).
\end{aligned}$$

The term with $\boldsymbol{\phi}^h$ in the last two arguments vanishes by (6.53). Now, all terms are estimated separately, using (6.36), Young's inequality (A.5), and the stability estimates (6.49) and (6.58). For the first term, one obtains

$$\begin{aligned}
n_{\text{skew}}\left(\boldsymbol{u}, \boldsymbol{\eta}, \boldsymbol{\phi}^h\right) &\le C \left\|\nabla \boldsymbol{u}\right\|_{L^2(\Omega)} \left\|\nabla \boldsymbol{\eta}\right\|_{L^2(\Omega)} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \\
&\le \frac{2C}{\nu} \left\|\nabla \boldsymbol{\eta}\right\|_{L^2(\Omega)}^2 \left\|\nabla \boldsymbol{u}\right\|_{L^2(\Omega)}^2 + \frac{\nu}{8} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 \\
&\le \frac{C}{\nu^3} \left\|\nabla \boldsymbol{\eta}\right\|_{L^2(\Omega)}^2 \left\|\boldsymbol{f}\right\|_{H^{-1}(\Omega)}^2 + \frac{\nu}{8} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2.
\end{aligned}$$

The estimate for the third term is performed analogously, yielding

$$n_{\text{skew}}\left(\boldsymbol{\eta}, \boldsymbol{u}^h, \boldsymbol{\phi}^h\right) \le \frac{C}{\nu^3} \left\|\nabla \boldsymbol{\eta}\right\|_{L^2(\Omega)}^2 \left\|\boldsymbol{f}\right\|_{H^{-1}(\Omega)}^2 + \frac{\nu}{8} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2.$$

The problematic term, for which the assumption on the smallness of the data is required, is the last one. With (6.57), (6.58), and (6.60), one gets

$$\begin{aligned}
n_{\text{skew}}\left(\boldsymbol{\phi}^h, \boldsymbol{u}^h, \boldsymbol{\phi}^h\right) &\le N_{\text{div}}^h \left\|\nabla \boldsymbol{u}^h\right\|_{L^2(\Omega)} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 \\
&\le \frac{4N_{\text{div}}^h \left\|\boldsymbol{f}\right\|_{H^{-1}(\Omega)}}{\nu^2} \left(\frac{\nu}{4} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2\right) \\
&\le \frac{\nu}{4} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2.
\end{aligned}$$

This term can be absorbed in the left-hand side of (6.63).

Substituting all estimates in (6.63) gives

$$\left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 \le C\left(\|\nabla\boldsymbol{\eta}\|_{L^2(\Omega)}^2 + \frac{1}{\nu^2}\left\|p-q^h\right\|_{L^2(\Omega)}^2 + \frac{1}{\nu^4}\|\nabla\boldsymbol{\eta}\|_{L^2(\Omega)}^2\|\boldsymbol{f}\|_{H^{-1}(\Omega)}^2\right).$$

The application of the triangle inequality finally leads to

$$\begin{aligned}\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)} &\le \|\nabla\boldsymbol{\eta}\|_{L^2(\Omega)} + \left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}\\ &\le C\left(\left(1+\frac{1}{\nu^2}\|\boldsymbol{f}\|_{H^{-1}(\Omega)}\right)\|\nabla\boldsymbol{\eta}\|_{L^2(\Omega)} + \frac{1}{\nu}\left\|p-q^h\right\|_{L^2(\Omega)}\right)\end{aligned}$$

for all $q^h \in Q^h$. This estimate gives the statement of the theorem. ■

Remark 6.29 (On Theorem 6.28)

- From the proof of Theorem 6.28, it is clear that condition (6.60) can be relaxed to

$$\frac{N_{\text{div}}^h\|\boldsymbol{f}\|_{H^{-1}(\Omega)}}{\nu^2} \le q < 1$$

  by applying different scalings in the applications of Young's inequality. However, large values of q in the analysis lead to a large constant C in the error bound (6.61).
- Instead of (6.64), the difference of the nonlinear terms can be rearranged as follows

$$\begin{aligned}&n_{\text{skew}}\left(\boldsymbol{u},\boldsymbol{u},\boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{\phi}^h\right)\\ &= n_{\text{skew}}\left(\boldsymbol{u},\boldsymbol{u},\boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\boldsymbol{u}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right) + n_{\text{skew}}\left(\boldsymbol{u}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right)\\ &\quad - n_{\text{skew}}\left(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{\phi}^h\right)\\ &= n_{\text{skew}}\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right) + n_{\text{skew}}\left(\boldsymbol{u}^h,\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{\phi}^h\right) \qquad (6.65)\\ &= n_{\text{skew}}\left(\boldsymbol{\eta},\boldsymbol{u},\boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\boldsymbol{\phi}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right) + n_{\text{skew}}\left(\boldsymbol{u}^h,\boldsymbol{\eta},\boldsymbol{\phi}^h\right)\\ &\quad - n_{\text{skew}}\left(\boldsymbol{u}^h,\boldsymbol{\phi}^h,\boldsymbol{\phi}^h\right).\end{aligned}$$

  The last term vanishes since the last both arguments are identical. Now, the problematic term is the second one. This term can be hidden in the left-hand side if a condition of the form

$$\frac{N_{\text{div}}\|\boldsymbol{f}\|_{H^{-1}(\Omega)}}{\nu^2} \le \frac{1}{4}$$

  is satisfied.

- Another rearrangement of the convective terms, which can be found in the literature, uses the definition (6.25) of the skew-symmetric form and (6.13), leading to

$$\begin{aligned} & n_{\text{skew}}\left(\boldsymbol{u},\boldsymbol{u},\boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{\phi}^h\right) \\ &= n_{\text{conv}}\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right) + n_{\text{conv}}\left(\boldsymbol{u}^h,\boldsymbol{\eta},\boldsymbol{\phi}^h\right) - \frac{1}{2}\left(\nabla\cdot\boldsymbol{u}^h,\boldsymbol{\phi}^h\cdot I^h\boldsymbol{u}\right). \end{aligned}$$

  Estimating these terms requires higher regularity assumptions, e.g., $\nabla\boldsymbol{u} \in L^\infty(\Omega)$, compare Arndt et al. (2015). □

Theorem 6.30 (Finite Element Error Estimate for the $L^2(\Omega)$ Norm of the Pressure) *Let the assumptions of Theorem 6.28 be fulfilled, then the following error estimate for the pressure holds*

$$\begin{aligned} \left\|p-p^h\right\|_{L^2(\Omega)} \le{}& C\frac{\nu}{\beta_{\text{is}}^h}\left(1+\frac{\|\boldsymbol{f}\|_{H^{-1}(\Omega)}}{\nu^2}\right)^2 \inf_{\boldsymbol{v}^h\in V_{\text{div}}^h}\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{v}^h\right)\right\|_{L^2(\Omega)} \\ &+C\left(1+\frac{1}{\beta_{\text{is}}^h}\left(1+\frac{\|\boldsymbol{f}\|_{H^{-1}(\Omega)}}{\nu^2}\right)\right)\inf_{q^h\in Q^h}\left\|p-q^h\right\|_{L^2(\Omega)}. \end{aligned} \tag{6.66}$$

The constants do not depend on the mesh size.

Proof The proof follows the lines of the proofs of Theorems 4.25 and 5.15.

The triangle inequality gives for all $q^h \in Q^h$

$$\left\|p-p^h\right\|_{L^2(\Omega)} \le \left\|p-q^h\right\|_{L^2(\Omega)} + \left\|p^h-q^h\right\|_{L^2(\Omega)}. \tag{6.67}$$

The estimate of the second term starts with the discrete inf-sup condition (3.51) and the insertion of the finite element problem (6.51) as well as the variational form of the steady-state Navier–Stokes equations (6.2)

$$\begin{aligned} &\left\|p^h-q^h\right\|_{L^2(\Omega)} \\ &\le \frac{1}{\beta_{\text{is}}^h}\sup_{\boldsymbol{v}^h\in V^h\setminus\{\boldsymbol{0}\}}\frac{b\left(\boldsymbol{v}^h,p^h-q^h\right)}{\|\nabla\boldsymbol{v}^h\|_{L^2(\Omega)}} \\ &= \frac{1}{\beta_{\text{is}}^h}\sup_{\boldsymbol{v}^h\in V^h\setminus\{\boldsymbol{0}\}}\left(\frac{\nu\left(\nabla(\boldsymbol{u}-\boldsymbol{u}^h),\nabla\boldsymbol{v}^h\right)-\left(\nabla\cdot\boldsymbol{v}^h,p-q^h\right)}{\|\nabla\boldsymbol{v}^h\|_{L^2(\Omega)}}\right. \\ &\qquad\left.+\frac{n_{\text{skew}}\left(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}^h\right)-n_{\text{skew}}\left(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{v}^h\right)}{\|\nabla\boldsymbol{v}^h\|_{L^2(\Omega)}}\right). \end{aligned} \tag{6.68}$$

In the next step, an identity like (6.64) is used

$$
\begin{aligned}
& n_{\text{skew}}\left(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}^h\right) - n_{\text{skew}}\left(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{v}^h\right) \\
&\quad = n_{\text{skew}}\left(\boldsymbol{u},\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{v}^h\right) + n_{\text{skew}}\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{v}^h\right).
\end{aligned}
$$

Inserting this identity, using the Cauchy–Schwarz inequality (A.10), the estimate (6.36) for the trilinear term, (3.170), and the stability estimates (6.49) and (6.58) leads to

$$
\begin{aligned}
& \left\|p^h - q^h\right\|_{L^2(\Omega)} \\
& \le \frac{1}{\beta_{\text{is}}^h}\Big(\nu\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)} + C\left\|\nabla\boldsymbol{u}\right\|_{L^2(\Omega)}\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)} \\
& \quad + C\left\|\nabla\boldsymbol{u}^h\right\|_{L^2(\Omega)}\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)} + \left\|p-q^h\right\|_{L^2(\Omega)}\Big) \\
& = \frac{C}{\beta_{\text{is}}^h}\left(\left(\nu + \frac{\|\boldsymbol{f}\|_{H^{-1}(\Omega)}}{\nu}\right)\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)} + \left\|p-q^h\right\|_{L^2(\Omega)}\right).
\end{aligned}
$$

Using this estimate in (6.67) gives

$$
\begin{aligned}
& \left\|p - p^h\right\|_{L^2(\Omega)} \\
& \le \frac{C}{\beta_{\text{is}}^h}\nu\left(1 + \frac{\|\boldsymbol{f}\|_{H^{-1}(\Omega)}}{\nu^2}\right)\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)} + \left(1 + \frac{C}{\beta_{\text{is}}^h}\right)\left\|p-q^h\right\|_{L^2(\Omega)}.
\end{aligned}
$$

Inserting now estimate (6.61) for $\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)}$ finishes the proof. ■

Remark 6.31 (Dual Linearized Problem) The application of the classical argument by Aubin (1967) and Nitsche (1968) for proving an estimate for the $L^2(\Omega)$ error of the velocity is based on the consideration of a dual linearized problem for the steady-state Navier–Stokes equations. The procedure for deriving the operator of the dual problem is explained in Remark 4.27 and the linearization of the Navier–Stokes equations will be discussed in Sect. 6.3.

Consider the Navier–Stokes equations in V_{div} and let $\boldsymbol{w} \in V_{\text{div}}$ be an arbitrary element. Then, the linearization of the convective term of the Navier–Stokes equations at $\boldsymbol{w}$ is given by the left-hand side of (6.86) below, which reads as follows

$$
n(\boldsymbol{w},\boldsymbol{u},\boldsymbol{v}) + n(\boldsymbol{u},\boldsymbol{w},\boldsymbol{v}) \quad \forall\ \boldsymbol{v} \in V_{\text{div}}. \tag{6.69}
$$

Here, $\boldsymbol{u} \in V_{\text{div}}$ is the solution of the Navier–Stokes equations.

The first term of (6.69) can be reformulated using (6.24)

$$
n(\boldsymbol{w},\boldsymbol{u},\boldsymbol{v}) = -n(\boldsymbol{w},\boldsymbol{v},\boldsymbol{u}).
$$

For the second term, (6.12) is applied, which gives

$$n(\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v}) = \left((\nabla \boldsymbol{w})^T \boldsymbol{v}, \boldsymbol{u}\right).$$

Denoting the velocity field of the dual linearized problems by $\boldsymbol{\phi}$ and applying the approach explained in Remark 4.27 reveals that one gets as contribution of the linearized convective terms to the dual equation

$$-n(\boldsymbol{w}, \boldsymbol{\phi}, \boldsymbol{v}) + \left((\nabla \boldsymbol{w})^T \boldsymbol{\phi}, \boldsymbol{v}\right) = -n(\boldsymbol{w}, \boldsymbol{\phi}, \boldsymbol{v}) - \left((\nabla \boldsymbol{\phi})^T \boldsymbol{w}, \boldsymbol{v}\right), \tag{6.70}$$

where (6.12) and (6.24) have been used for deriving the second form. The terms from (6.70) correspond to the strong forms

$$-(\boldsymbol{w} \cdot \nabla)\boldsymbol{\phi} + (\nabla \boldsymbol{w})^T \boldsymbol{\phi} \quad \text{or} \quad -(\boldsymbol{w} \cdot \nabla)\boldsymbol{\phi} - (\nabla \boldsymbol{\phi})^T \boldsymbol{w}.$$

These are a convective term and a reactive term. However the convection field of the dual problem (6.70) is the negative of the convection field of the linearized Navier–Stokes problem (6.69). Note that with (6.12) and (6.24), the terms in (6.70) can be written in the form

$$-n(\boldsymbol{w}, \boldsymbol{\phi}, \boldsymbol{v}) + \left((\nabla \boldsymbol{w})^T \boldsymbol{\phi}, \boldsymbol{v}\right) = n(\boldsymbol{w}, \boldsymbol{v}, \boldsymbol{\phi}) + n(\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{\phi}). \tag{6.71}$$

Deriving the other terms for the dual linearized problem in the same way as for the Stokes equations, compare Remark 4.27, and using the skew-symmetric form of the convective term in (6.71), such that this form can be used in the finite element error analysis, leads to the following dual linearized problem: For given $\hat{\boldsymbol{f}} \in L^2(\Omega)$ and $\boldsymbol{w} \in V_{\text{div}}$, find $(\boldsymbol{\phi}_{\hat{f}}, \xi_{\hat{f}}) \in V \times Q$ such that

$$\begin{aligned}(\nu \nabla \boldsymbol{v}, \nabla \boldsymbol{\phi}_{\hat{f}}) + n_{\text{skew}}\left(\boldsymbol{w}, \boldsymbol{v}, \boldsymbol{\phi}_{\hat{f}}\right) + n_{\text{skew}}\left(\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{\phi}_{\hat{f}}\right) + (\nabla \cdot \boldsymbol{v}, \xi_{\hat{f}}) &= \left(\hat{\boldsymbol{f}}, \boldsymbol{v}\right), \\ (\nabla \cdot \boldsymbol{\phi}_{\hat{f}}, q) &= 0 \end{aligned} \tag{6.72}$$

for all $(\boldsymbol{v}, q) \in V \times Q$. Again, the general dual linearized problem might have a non-vanishing right-hand side in the second equation. Problem (6.72) is called regular, if the mapping

$$\left(\boldsymbol{\phi}_{\hat{f}}, \xi_{\hat{f}}\right) \mapsto \hat{\boldsymbol{f}} \tag{6.73}$$

is an isomorphism from $\left(H^2(\Omega) \cap V\right) \times \left(H^1(\Omega) \cap Q\right)$ onto $L^2(\Omega)$, see Remark 4.27 for a short discussion on the regularity of such mappings. □

Theorem 6.32 (Finite Element Error Estimate for the $L^2(\Omega)$ Norm of the Velocity) *Let the assumption of Theorem 6.28 hold and let* $(\boldsymbol{\phi}_{\hat{f}}, \xi_{\hat{f}})$ *be the solution* *of* (6.72). *Then, it holds for the* $L^2(\Omega)$ *error of the velocity*

$$
\begin{aligned}
\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{L^2(\Omega)} \leq & \sup_{\hat{\boldsymbol{f}}\in L^2(\Omega)\setminus\{\boldsymbol{0}\}} \frac{1}{\left\|\hat{\boldsymbol{f}}\right\|_{L^2(\Omega)}} \\
& \times\left\{ \inf_{\boldsymbol{\phi}^h\in V_{\text{div}}^h} \left[\left(\nu \left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)} + \left\|p-q^h\right\|_{L^2(\Omega)}\right.\right.\right. \\
& \left. + C\left\|\nabla \boldsymbol{u}\right\|_{L^2(\Omega)} \left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}\right) \left\|\nabla\left(\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}-\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)} \\
& \left. + C\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}^2 \left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}\right] \\
& \left. + \inf_{r^h\in Q^h} \left\|\xi_{\hat{\boldsymbol{f}}} - r^h\right\|_{L^2(\Omega)} \left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}\right\}.
\end{aligned}
\tag{6.74}
$$

Proof The proof proceeds analogously to the proof of Theorem 4.28. The only difference is the appearance of the trilinear terms in the dual problem (6.72), which is considered for $\boldsymbol{w}=\boldsymbol{u}$, and in introducing these expressions on the right-hand side of (4.32). Considering only these terms, using the test function $\boldsymbol{v}=\boldsymbol{u}-\boldsymbol{u}^h$, and applying (6.64) yields

$$
\begin{aligned}
& n_{\text{skew}}\left(\boldsymbol{u},\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}\right) + n_{\text{skew}}\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{u},\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}\right) + n_{\text{skew}}\left(\boldsymbol{u},\boldsymbol{u},\boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{\phi}^h\right) \\
&= n_{\text{skew}}\left(\boldsymbol{u},\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}\right) + n_{\text{skew}}\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{u},\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}\right) + n_{\text{skew}}\left(\boldsymbol{u},\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{\phi}^h\right) \\
&\quad + n_{\text{skew}}\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{\phi}^h\right) \\
&= n_{\text{skew}}\left(\boldsymbol{u},\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}-\boldsymbol{\phi}^h\right) + n_{\text{skew}}\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{u},\boldsymbol{\phi}_{\hat{\boldsymbol{f}}}-\boldsymbol{\phi}^h\right) \\
&\quad + n_{\text{skew}}\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{\phi}^h\right).
\end{aligned}
$$

All these terms can be estimated with (6.36). The other terms are the same terms as for the Stokes problem and they are estimated as shown in (4.34). Summarizing all estimates gives the error estimate (6.74). ■

Corollary 6.33 (Finite Element Error Estimates for Conforming Pairs of Finite Element Spaces) *Let the assumptions of Theorem 6.28 be fulfilled, let* $\Omega\subset\mathbb{R}^d$, $d\in\{2,3\}$, *be a bounded domain with polyhedral and Lipschitz continuous boundary which is decomposed by a regular and quasi-uniform family of triangulations* $\{\mathcal{T}^h\}$, *and let* $(\boldsymbol{u},p)$ *be the unique solution of the steady-state Navier–Stokes equations* (6.2) *with* $\boldsymbol{u}\in H^{k+1}(\Omega)\cap V$ *and* $p\in H^k(\Omega)\cap Q$. *Then for the inf-sup* *stable pairs of finite element spaces*

- P_k^{bubble}/P_k, $k = 1$ *(MINI element),*
- P_k/P_{k-1}, Q_k/Q_{k-1}, $k \geq 2$ *(Taylor–Hood element),*
- $P_2^{\text{bubble}}/P_1^{\text{disc}}$, $P_3^{\text{bubble}}/P_2^{\text{disc}}$, $P_2^{\text{BR}}/P_1^{\text{disc}}$, $Q_k/P_{k-1}^{\text{disc}}$, $k \geq 2$,

the following error estimates hold

$$\left\|\nabla(\boldsymbol{u} - \boldsymbol{u}^h)\right\|_{L^2(\Omega)} \leq Ch^k \left(\frac{1}{\nu^2}\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \frac{1}{\nu}\|p\|_{H^k(\Omega)}\right), \tag{6.75}$$

$$\left\|p - p^h\right\|_{L^2(\Omega)} \leq Ch^k \left(\frac{1}{\nu^3}\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \frac{1}{\nu^2}\|p\|_{H^k(\Omega)}\right). \tag{6.76}$$

The constant in (6.75) *depends either on the inverse of the discrete inf-sup constant* β_{is}^h *or on the inverse of local inf-sup constants, compare Remark 4.29. The constant in* (6.76) *does depend on the inverse of the discrete inf-sup constant* β_{is}^h.

If the solution $(\boldsymbol{\phi}_{\hat{f}}, \xi_{\hat{f}})$ *of the dual linearized Navier–Stokes equations is regular in the sense of Remark 6.31, then it holds additionally*

$$\left\|\boldsymbol{u} - \boldsymbol{u}^h\right\|_{L^2(\Omega)} \leq Ch^{k+1}\left(\|\boldsymbol{u}\|_{H^{k+1}(\Omega)} + \|p\|_{H^k(\Omega)}\right). \tag{6.77}$$

The constant C depends on the constants of estimates (6.75), (6.76)*, and on inverse powers of* ν.

Proof The first two estimates follow directly from (6.61), (6.66), and the approximation properties of the finite element spaces. If the constant in (6.61) depends on local inf-sup constants or the global discrete inf-sup constant depends on whether (3.71) can be applied or (3.65) has to be used. The dependency of the constant in (6.76) on $\left(\beta_{\text{is}}^h\right)^{-1}$ follows directly from (6.66).

For deriving estimate (6.77), one uses estimate (6.74) and applies (6.75) and (6.76). The additional power of h is a result of the regularity of the dual linearized Navier–Stokes equations and the application of the interpolation estimate, see Theorem C.13, compare the proof of Corallary 4.30. Note that the term $\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}$ in (6.74) is estimated in (6.49) by the right-hand side of the Navier–Stokes equations. The other term can be bounded using the triangle inequality, the interpolation estimate (C.7), and the regularity of the dual linearized problem by

$$\left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \leq \left\|\nabla\boldsymbol{\phi}_{\hat{f}}\right\|_{L^2(\Omega)} + \left\|\nabla\left(\boldsymbol{\phi}_{\hat{f}} - \boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)} \leq C\left\|\hat{\boldsymbol{f}}\right\|_{L^2(\Omega)}.$$

■

Remark 6.34 (Dependency on ν *in Comparison with the Oseen Equations)* It shall be recalled that the case of very small ν is not of practical interest for the steady-state Navier–Stokes equations, compare Remark 6.16.

For the Oseen equations, the constants in the error bounds for the $L^2(\Omega)$ errors of the gradient of the velocity and the pressure depend only on ν^{-1}, compare the discussion in Remark 5.17. It follows from (6.75) and (6.66) that there is a stronger dependency for the Navier–Stokes equations. Formally, the difference arises on the one hand from absorbing the term $\|\boldsymbol{b}\|_{L^\infty(\Omega)}$ from the convection field in the constants of the error bounds for the Oseen equations because $\|\boldsymbol{b}\|_{L^\infty(\Omega)} = \mathcal{O}(1)$ is assumed. On the other hand, the terms $\|\nabla \boldsymbol{u}\|_{L^2(\Omega)}$ and $\left\|\nabla \boldsymbol{u}^h\right\|_{L^2(\Omega)}$ are estimated with the stability bounds (6.49) and (6.58) in the case of the Navier–Stokes equations, which gives an extra negative power of ν. Assuming that the convection field of the Oseen equations behaves like the convection field of the Navier–Stokes equations, then the constant C_{os} from (5.16) scales like $\mathcal{O}\left(\nu^{-3/2}\right)$ such that one obtains in the error bounds (5.23) and (5.24) (for $c = 0$) the same dependency on the viscosity as for the Navier–Stokes equations. □

Example 6.35 (Analytical Example Which Supports the Error Estimates) Example D.3 is considered, for detailed information to the simulations it is also referred to Example 4.31.

Figures 6.1, 6.2, and 6.3 show results that were obtained for $\nu = 10^{-2}$ and the grids from Fig. 4.2. The stopping criterion for the solution of the nonlinear problems was the requirement that the Euclidean norm of the residual vector was less than 10^{-10}. The convective form of the convective term was used in the simulations.

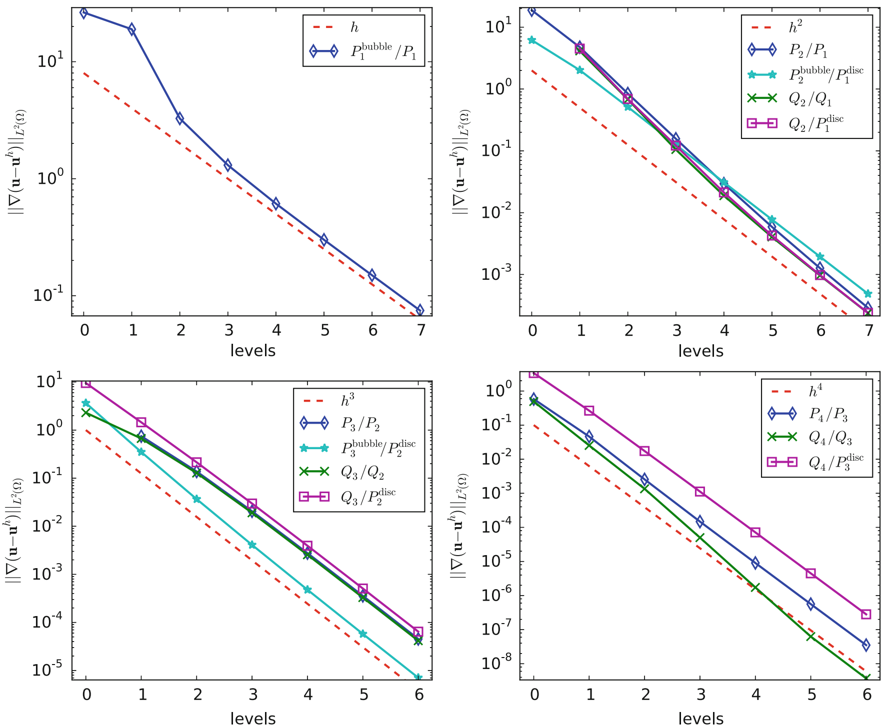

Fig. 6.1 Example 6.35. Convergence of the errors $\left\|\nabla(\boldsymbol{u} - \boldsymbol{u}^h)\right\|_{L^2(\Omega)}$ for different discretizations with different orders k

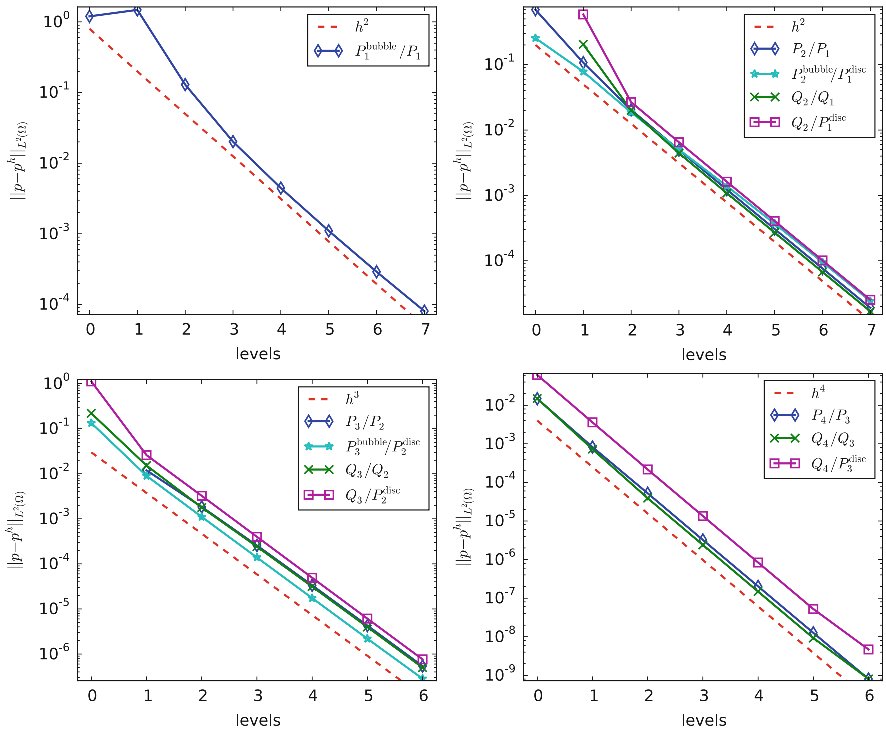

Fig. 6.2 Example 6.35. Convergence of the errors $\|p - p^h\|_{L^2(\Omega)}$ for different discretizations with different orders k

The order of convergence for errors in different norms coincide generally with the predictions from the numerical analysis. Only for the $L^2(\Omega)$ norm of the pressure and the MINI element $P_1^{\mathrm{bubble}}/P_1$, a higher order than expected can be observed. There are generally little differences with respect to the accuracy among different pairs of spaces with the same order. Only the results obtained with $P_3^{\mathrm{bubble}}/P_2^{\mathrm{disc}}$ are somewhat more accurate than the results of the other pairs with third order velocity and second order pressure. For some pairs, the nonlinear problem on the coarsest grid could not be solved. Round-off errors, due to the choice of the stopping criterion, can be seen for the highest order pairs on the finest meshes.

Figure 6.4 presents results for the Q_2/Q_1 finite element and different values of ν. One can observe on coarser grids larger velocity errors for smaller coefficients ν. For finer grids, the errors $\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}$ and $\left\|p-p^h\right\|_{L^2(\Omega)}$ do not show a dependency on ν. Also the curve for $\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{L^2(\Omega)}$ and $\nu = 1/400$ seems to converge to the other curves. □

Example 6.36 (Flow Around a Cylinder in Two Dimensions) This problem is described in Example D.5.

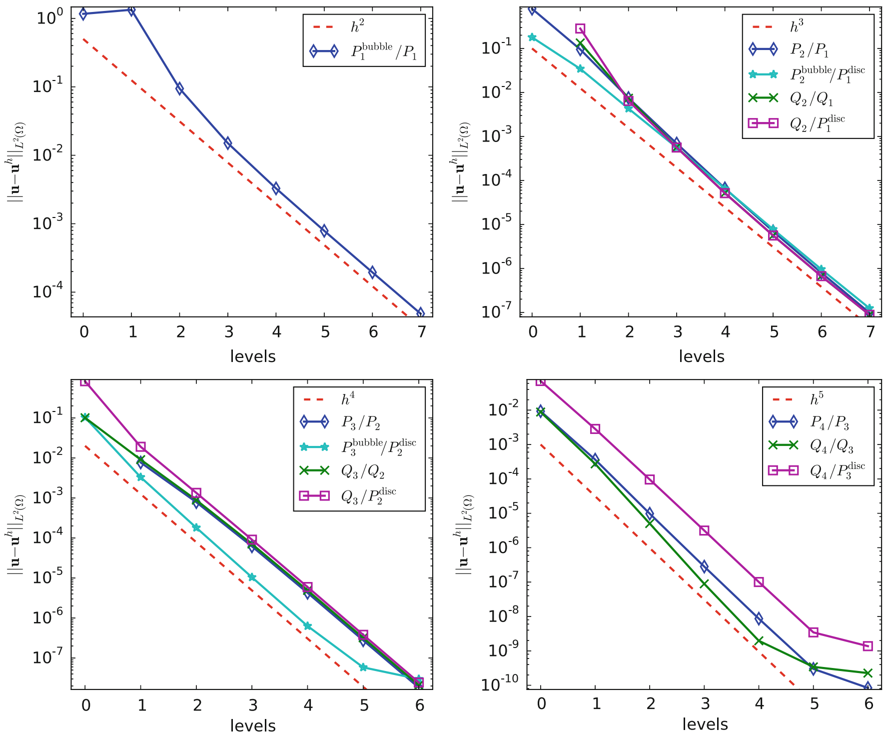

Fig. 6.3 Example 6.35. Convergence of the errors $\|\boldsymbol{u}-\boldsymbol{u}^h\|_{L^2(\Omega)}$ for different discretizations with different orders k

Simulations were performed with the convective form $n_{\rm conv}(\cdot,\cdot,\cdot)$ of the convective term and different inf-sup stable pairs of finite element spaces on triangular and quadrilateral grids. The initial grids (level 0) are presented in Fig. 6.5. The considered problem, with Dirichlet conditions at the outflow (D.14), was studied comprehensively in John and Matthies (2001). In these studies, it was shown that the use of isoparametric finite elements at the cylinder is essential for obtaining accurate results for higher order finite element discretizations. For isoparametric elements, the same functions are used for the construction of the finite element spaces and the definition of the map from the reference cell to the physical mesh cell. In this way, one obtains a better approximation of the boundary $\Gamma_{\rm cyl}$, but not yet the correct representation. Isoparametric finite elements were used in the studies presented here. The approximation of the boundary of the cylinder is denoted by $\Gamma_{\rm cyl}^h$. The solution of the nonlinear systems was stopped if the Euclidean norm of the residual vector was less than 10^{-15}.

For computing the drag and the lift coefficient, the volume formulations (D.16) and (D.17) with the convective form of the convection term were used. In these formulations, one has to specify the functions $\boldsymbol{w}_d$ and $\boldsymbol{w}_l$ in the interior of Ω.

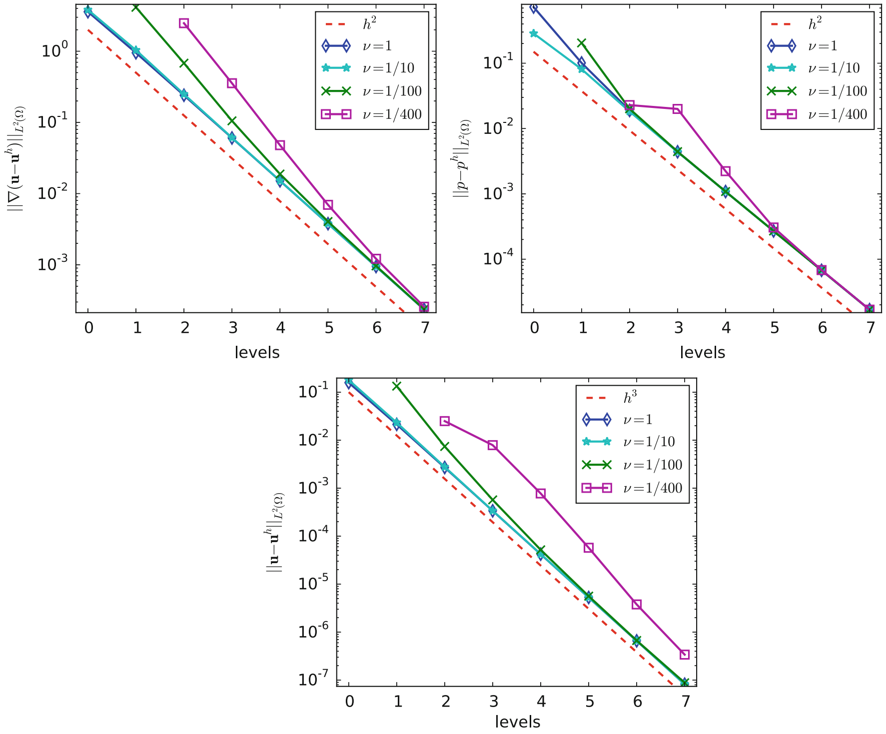

Fig. 6.4 Example 6.35. Convergence of the errors for different values of ν, Q_2/Q_1 finite element

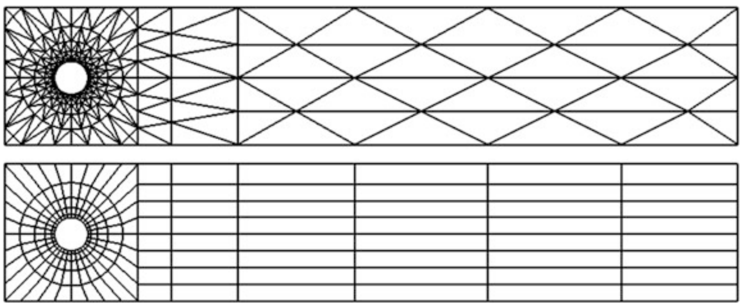

Fig. 6.5 Example 6.36. Initial grids (level 0) used for the simulations

Because these functions are up to the boundary arbitrary functions, one can use in actual computations finite element functions with appropriate boundary values. For the results presented below, the functions were chosen in such a way that they have the same order as the finite element velocity, the degrees of freedom at Γ_{cyl}^h were set to be one in the needed component, and all other degrees of freedom were set to be

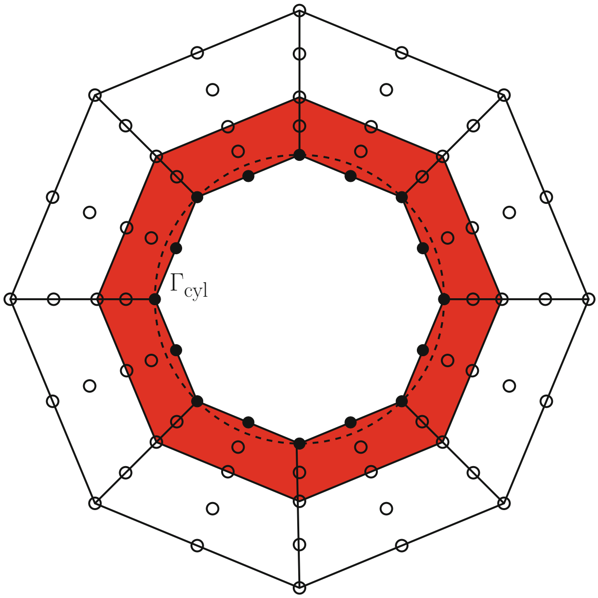

Fig. 6.6 Example 6.36. Choice of $\boldsymbol{w}_d$ and $\boldsymbol{w}_l$ for Q_2. In the *filled bullets*, the value 1 was set in the respective component, in the *empty bullets* the finite element functions have value $\mathbf{0}$. The *filled mesh cells* form the domain of integration in (D.16) and (D.17)

zero, see Fig. 6.6. With this approach, only the evaluation of volume integrals in one layer of mesh cells around the circle is necessary for computing the coefficients.

The used grids possess nodes in the points $(0.15, 0.2)$ and $(0.25, 0.2)$. Thus, the finite element pressure for discretizations with discontinuous pressure approximation is generally not continuous in these points. For computing the difference of the pressure (D.18), the values of the finite element pressure coming from all mesh cells with the node $(0.15, 0.2)$ or $(0.25, 0.2)$, respectively, were averaged.

Results for the drag coefficient, the lift coefficient, and the difference of the pressure between the front and the back of the cylinder are presented in Figs. 6.7, 6.8, 6.9, and 6.10. It can be seen that many results show a certain order of convergence. To the best of our knowledge, a numerical analysis of this phenomenon is not available. A possible approach was presented in John et al. (1998). It is however not clear if the regularity assumptions on the solution of the continuous problem assumed in this paper are always true.

Comparing the results of discretizations with different order, the higher accuracy of third order velocity/second order pressure compared with second order velocity/first order pressure can be observed also for quantities of interest that are not errors in norms of Sobolev spaces. The comparatively inaccurate results for the $P_2^{\text{bubble}}/P_1^{\text{disc}}$ pair of spaces are notable. It can be also seen that with discontinuous pressure approximations and averaging of pressure values, comparatively inaccurate results for the pressure difference are obtained, see Fig. 6.10. □

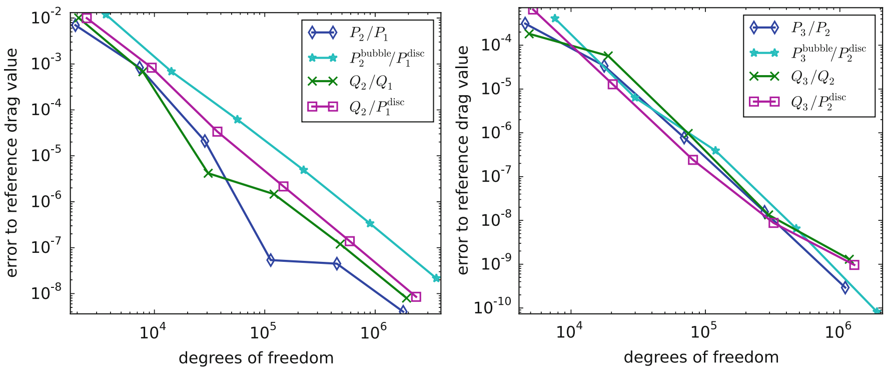

Fig. 6.7 Example 6.36. Convergence of the drag coefficient for different pairs of finite element spaces, boundary conditions (D.15), reference value (D.19). The results for boundary condition (D.14) are almost identical

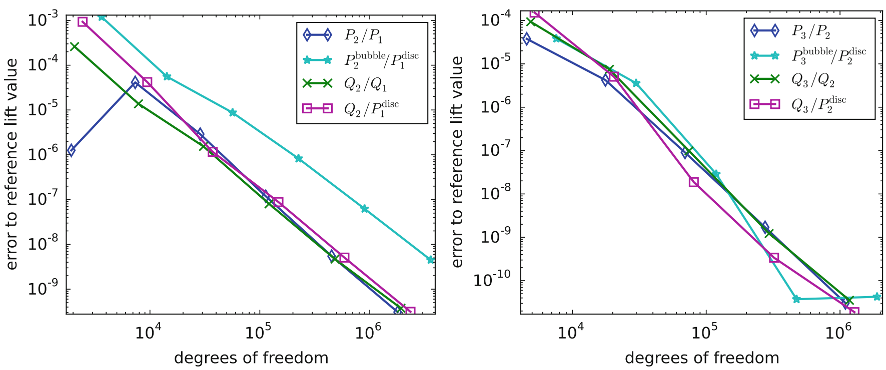

Fig. 6.8 Example 6.36. Convergence of the lift coefficient for different pairs of finite element spaces, Dirichlet boundary conditions (D.14), reference value (D.20)

Remark 6.37 (Convective Forms in Simulations) Despite the lack of a finite element error analysis, the convective form of the convective term is often used in simulations. But also the use of the other forms can be found in the literature. In particular, the rotational form became somewhat popular in recent years.

Comprehensive studies on the advantages and drawbacks of the different forms of the convective term were performed in Rockel (2013). Concerning the accuracy of the results, generally only minor differences could be observed between using the convective, the skew-symmetric, and the divergence form of the convective term. The rotational form led sometimes to notably more inaccurate numerical solutions. It is reported in Olshanskii (2002) that applying the grad-div stabilization in this case improves the accuracy of the results. □

Remark 6.38 (Other Discretizations) All methods that were described for the Stokes and Oseen equations, like stabilizations with respect to the violation of the

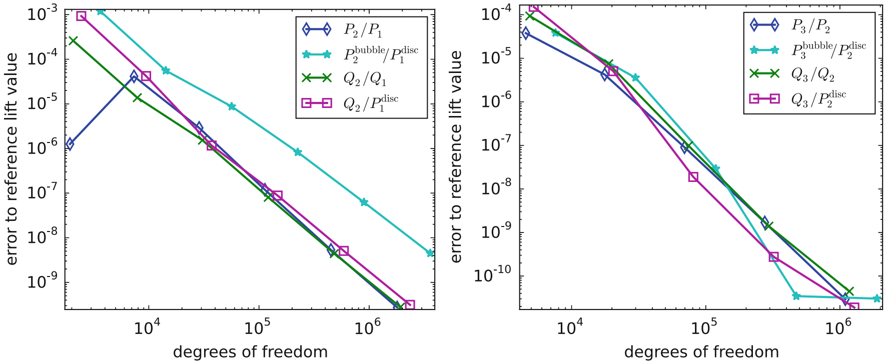

Fig. 6.9 Example 6.36. Convergence of the lift coefficient for different pairs of finite element spaces, do-nothing boundary conditions (D.15), reference value (D.22)

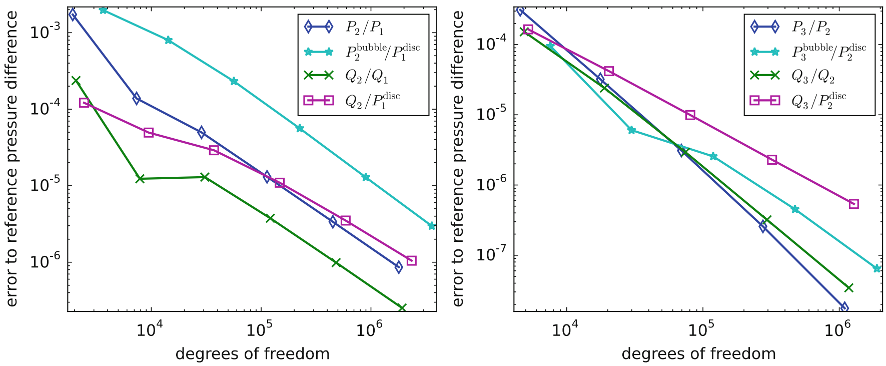

Fig. 6.10 Example 6.36. Convergence of the pressure difference for different pairs of finite element spaces, boundary conditions (D.15), reference value (D.21). The results for boundary condition (D.14) are almost identical

discrete inf-sup condition, stabilizations with respect to dominating convection, and non-conforming finite element spaces, can be applied in the numerical simulation of the steady-state Navier–Stokes equations. The detailed discussion of all these methods is beyond the scope of this monograph, therefore it will be referred to the literature. □

6.3 Iteration Schemes for Solving the Nonlinear Problem

Remark 6.39 (General Fixed Point Iteration) The Navier–Stokes equations (6.3) can be written in operator form, see also Remark 3.4 for this concept,

$$\mathbf{0} = \boldsymbol{f} - A\boldsymbol{u} - N_{\boldsymbol{u}}\boldsymbol{u} - B'p + B\boldsymbol{u} \quad \text{in } V' \times Q',$$

where $N_{\boldsymbol{u}} : V \to V'$ is the operator for the nonlinear convective term. Applying now an injective linear operator $N_{\text{lin}}^{-1} : V' \times Q' \to V \times Q$ yields

$$\mathbf{0} = N_{\text{lin}}^{-1}\mathbf{0} = \boldsymbol{N}_{\text{lin}}^{-1}\left(\boldsymbol{f} - A\boldsymbol{u} - N_{\boldsymbol{u}}\boldsymbol{u} - B'p + B\boldsymbol{u}\right) \quad \text{in } V \times Q.$$

Then, the operator $\boldsymbol{N}_{\text{lin}}^{-1}\,(\boldsymbol{f} - A\boldsymbol{u} - N_{\boldsymbol{u}}\boldsymbol{u} - B'p + B\boldsymbol{u})$ is a map from $V \times Q \to V \times Q$ and to this map, the standard approach for a fixed point iteration can be applied: Given $\left(\boldsymbol{u}^{(m)}, p^{(m)}\right) \in V \times Q$, compute $\left(\boldsymbol{u}^{(m+1)}, p^{(m+1)}\right) \in V \times Q$ by

$$\begin{pmatrix} \boldsymbol{u}^{(m+1)} \\ p^{(m+1)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{u}^{(m)} \\ p^{(m)} \end{pmatrix} - \vartheta \boldsymbol{N}_{\text{lin}}^{-1} \begin{pmatrix} \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{V',V} - \boldsymbol{N}\left(\boldsymbol{u}^{(m)}; \boldsymbol{u}^{(m)}, p^{(m)}\right) \\ 0 \end{pmatrix}, \tag{6.78}$$

where

$$\boldsymbol{N}\left(\boldsymbol{w}; \boldsymbol{u}, p\right) = \begin{pmatrix} a(\boldsymbol{u}, \boldsymbol{v}) + n(\boldsymbol{w}, \boldsymbol{u}, \boldsymbol{v}) + b(\boldsymbol{v}, p) \\ b(\boldsymbol{u}, q) \end{pmatrix}$$

and $\vartheta \in (0, 1]$ is a damping factor. For convenience of notation, the operators are replaced by the bilinear forms.

An iteration of form (6.78) requires the solution of a linear problem. Let $\boldsymbol{N}_{\text{lin}} : \text{range}\left(\boldsymbol{N}_{\text{lin}}^{-1}\right) \to V' \times Q'$ be the inverse operator of $\boldsymbol{N}_{\text{lin}}^{-1}$, then the linear problem has the form

$$\boldsymbol{N}_{\text{lin}} \begin{pmatrix} \delta\boldsymbol{u}^{(m+1)} \\ \delta p^{(m+1)} \end{pmatrix} = \begin{pmatrix} \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{V',V} - \boldsymbol{N}\left(\boldsymbol{u}^{(m)}; \boldsymbol{u}^{(m)}, p^{(m)}\right) \\ 0 \end{pmatrix}.$$

Writing the update in the form

$$\begin{pmatrix} \delta\boldsymbol{u}^{(m+1)} \\ \delta p^{(m+1)} \end{pmatrix} = \begin{pmatrix} \tilde{\boldsymbol{u}}^{(m+1)} - \boldsymbol{u}^{(m)} \\ \tilde{p}^{(m+1)} - p^{(m)} \end{pmatrix},$$

the linear system can be reformulated for a new velocity and pressure solution

$$\boldsymbol{N}_{\text{lin}} \begin{pmatrix} \tilde{\boldsymbol{u}}^{(m+1)} \\ \tilde{p}^{(m+1)} \end{pmatrix} = \begin{pmatrix} \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{V',V} - \boldsymbol{N}\left(\boldsymbol{u}^{(m)}; \boldsymbol{u}^{(m)}, p^{(m)}\right) \\ 0 \end{pmatrix} + \boldsymbol{N}_{\text{lin}} \begin{pmatrix} \boldsymbol{u}^{(m)} \\ p^{(m)} \end{pmatrix}. \tag{6.79}$$

□

Remark 6.40 (Fixed Point Iteration with Scaled Stokes Equations) A simple iterative approach is obtained by setting

$$\boldsymbol{N}_{\text{lin}} = \boldsymbol{N}\left(\mathbf{0}; \tilde{\boldsymbol{u}}^{(m+1)}, \tilde{p}^{(m+1)}\right).$$

Inserting this expression in (6.79) gives

$$\begin{pmatrix} a(\tilde{\boldsymbol{u}}^{(m+1)}, \boldsymbol{v}) + b(\boldsymbol{v}, \tilde{p}^{(m+1)}) \\ b(\tilde{\boldsymbol{u}}^{(m+1)}, q) \end{pmatrix}$$

$$= \begin{pmatrix} \langle \boldsymbol{f}, \boldsymbol{v}\rangle_{V',V} - a(\boldsymbol{u}^{(m)}, \boldsymbol{v}) - n(\boldsymbol{u}^{(m)}, \boldsymbol{u}^{(m)}, \boldsymbol{v}) - b(\boldsymbol{v}, p^{(m)}) \\ -b(\boldsymbol{u}^{(m)}, q) \end{pmatrix}$$

$$+ \begin{pmatrix} a(\boldsymbol{u}^{(m)}, \boldsymbol{v}) + b(\boldsymbol{v}, p^{(m)}) \\ b(\boldsymbol{u}^{(m)}, q) \end{pmatrix}$$

$$= \begin{pmatrix} \langle \boldsymbol{f}, \boldsymbol{v}\rangle_{V',V} - n(\boldsymbol{u}^{(m)}, \boldsymbol{u}^{(m)}, \boldsymbol{v}) \\ 0 \end{pmatrix}.$$

These equations are scaled Stokes equation, see (4.153).

This approach requires only the solution of a scaled Stokes problem with the same matrix and with a different right-hand side in each iteration step. It is well known for poor convergence properties in the case that ν is not sufficiently large. That means, it converges only with the application of strong damping or there is even no convergence at all if the initial iterate is not sufficiently close to the solution. Therefore, this type of fixed point iteration is in general not recommended and it will not be considered further here. □

Remark 6.41 (Picard Iteration) The so-called Picard iteration is obtained by setting

$$\boldsymbol{N}_{\text{lin}} = \boldsymbol{N}\left(\boldsymbol{u}^{(m)}; \tilde{\boldsymbol{u}}^{(m+1)}, \tilde{p}^{(m+1)}\right).$$

One obtains for the linear system (6.79) to be solved

$$\begin{pmatrix} a(\tilde{\boldsymbol{u}}^{(m+1)}, \boldsymbol{v}) + n(\boldsymbol{u}^{(m)}, \tilde{\boldsymbol{u}}^{(m+1)}, \boldsymbol{v}) + b(\boldsymbol{v}, \tilde{p}^{(m+1)}) \\ b(\tilde{\boldsymbol{u}}^{(m+1)}, q) \end{pmatrix}$$

$$= \begin{pmatrix} \langle \boldsymbol{f}, \boldsymbol{v}\rangle_{V',V} - a(\boldsymbol{u}^{(m)}, \boldsymbol{v}) - n(\boldsymbol{u}^{(m)}, \boldsymbol{u}^{(m)}, \boldsymbol{v}) - b(\boldsymbol{v}, p^{(m)}) \\ -b(\boldsymbol{u}^{(m)}, q) \end{pmatrix}$$

$$+ \begin{pmatrix} a(\boldsymbol{u}^{(m)}, \boldsymbol{v}) + n(\boldsymbol{u}^{(m)}, \boldsymbol{u}^{(m)}, \boldsymbol{v}) + b(\boldsymbol{v}, p^{(m)}) \\ b(\boldsymbol{u}^{(m)}, q) \end{pmatrix}$$

$$= \begin{pmatrix} \langle \boldsymbol{f}, \boldsymbol{v}\rangle_{V',V} \\ 0 \end{pmatrix}. \tag{6.80}$$

The different forms of the convective terms on the left-hand side of (6.80) look as follows:

$$\begin{aligned}
n_{\text{conv}}\left(\boldsymbol{u}^{(m)}, \tilde{\boldsymbol{u}}^{(m+1)}, \boldsymbol{v}\right) &= \left(\left(\boldsymbol{u}^{(m)} \cdot \nabla\right) \tilde{\boldsymbol{u}}^{(m+1)}, \boldsymbol{v}\right), \\
n_{\text{div}}\left(\boldsymbol{u}^{(m)}, \tilde{\boldsymbol{u}}^{(m+1)}, \boldsymbol{v}\right) &= \left(\left(\boldsymbol{u}^{(m)} \cdot \nabla\right) \tilde{\boldsymbol{u}}^{(m+1)}, \boldsymbol{v}\right) + \frac{1}{2}\left(\left(\nabla \cdot \boldsymbol{u}^{(m)}\right) \tilde{\boldsymbol{u}}^{(m+1)}, \boldsymbol{v}\right), \\
n_{\text{rot}}\left(\boldsymbol{u}^{(m)}, \tilde{\boldsymbol{u}}^{(m+1)}, \boldsymbol{v}\right) &= \left(\left(\nabla \times \boldsymbol{u}^{(m)}\right) \times \tilde{\boldsymbol{u}}^{(m+1)}, \boldsymbol{v}\right), \\
n_{\text{skew}}\left(\boldsymbol{u}^{(m)}, \tilde{\boldsymbol{u}}^{(m+1)}, \boldsymbol{v}\right) &= \frac{1}{2}\left[\left(\left(\boldsymbol{u}^{(m)} \cdot \nabla\right) \tilde{\boldsymbol{u}}^{(m+1)}, \boldsymbol{v}\right) - \left(\left(\boldsymbol{u}^{(m)} \cdot \nabla\right) \boldsymbol{v}, \tilde{\boldsymbol{u}}^{(m+1)}\right)\right].
\end{aligned}$$

Thus, using $n_{\text{conv}}\left(\boldsymbol{u}^{(m)}, \tilde{\boldsymbol{u}}^{(m+1)}, \boldsymbol{v}\right)$, one has to solve in (6.80) Oseen equations with $\boldsymbol{b} = \boldsymbol{u}^{(m)}$ and $c = 0$. For $n_{\text{div}}\left(\boldsymbol{u}^{(m)}, \tilde{\boldsymbol{u}}^{(m+1)}, \boldsymbol{v}\right)$, one obtains Oseen equations with $\boldsymbol{b} = \boldsymbol{u}^{(m)}$ and $c = \nabla \cdot \boldsymbol{u}^{(m)}$. In both cases, the equations are dominated by convection if ν is small compared with $\left\|\boldsymbol{u}^{(m)}\right\|_{L^\infty(\Omega)}$. The use of $n_{\text{rot}}\left(\boldsymbol{u}^{(m)}, \tilde{\boldsymbol{u}}^{(m+1)}, \boldsymbol{v}\right)$ leads to Oseen equations with $\boldsymbol{b} = \boldsymbol{0}$ and the matrix $\left(\nabla \times \boldsymbol{u}^{(m)}\right)$ as coefficient of the reactive term. In this case, there is no convection term in the linear problem.

Note that, in all cases, for finite element functions, the assumptions on the coefficients of the Oseen equations which were made in the analysis of the Oseen equations, see Remark 5.2, are generally not satisfied by the coefficients coming from the fixed point iteration of the steady-state Navier–Stokes equations. That means, $\boldsymbol{u}^{h,(m)}$ is generally not weakly divergence-free, $\nabla \cdot \boldsymbol{u}^{h,(m)}$ might be negative, and numerical analysis for a matrix coefficients $\left(\nabla \times \boldsymbol{u}^{h,(m)}\right)$ is even not available in the literature. However, even if such an analysis can be performed, the natural extension of the assumptions on the scalar reaction coefficient to a matrix coefficient would be that the matrix is symmetric and positive semi-definite. These assumptions are generally not fulfilled by $\left(\nabla \times \boldsymbol{u}^{h,(m)}\right)$.

The skew-symmetric form of the convective term does not lead to an equation of Oseen type, since in the zeroth order term with respect to $\tilde{\boldsymbol{u}}^{(m+1)}$, the gradient of the test function appears instead of the test function itself.

Approximation $\boldsymbol{u}^{(m)}$, it changes in every iteration. □

Remark 6.42 (Picard Method: Structure of Matrices and Memory Requirements) After having applied an inf-sup stable finite element method, the linear system (6.80) has the saddle point form

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{0} \end{pmatrix}. \tag{6.81}$$

- *Convective form of the convective term.* In the finite element equation (6.80) that arises in the fixed point iteration for solving the nonlinearity, the term

$$n_{\mathrm{conv}}\left(\boldsymbol{u}^{h,(m)}, \tilde{\boldsymbol{u}}^{h,(m+1)}, \boldsymbol{v}^h\right) = \left(\left(\boldsymbol{u}^{h,(m)} \cdot \nabla\right) \tilde{\boldsymbol{u}}^{h,(m+1)}, \boldsymbol{v}^h\right) \tag{6.82}$$

appears on the left-hand side. Here, $\boldsymbol{u}^{h,(m)}$ is a known finite element function. Using the ansatz (4.89) for $\tilde{\boldsymbol{u}}^{h,(m+1)}$ and considering a test function $\boldsymbol{\phi}_i^h$, one gets for the convective form of the convective term

$$\left(\left(\boldsymbol{u}^{h,(m)} \cdot \nabla\right) \tilde{\boldsymbol{u}}^{h,(m+1)}, \boldsymbol{\phi}_i^h\right) = \sum_{j=1}^{3N_v} u_j^h \left(\left(\boldsymbol{u}^{h,(m)} \cdot \nabla\right) \boldsymbol{\phi}_j^h, \boldsymbol{\phi}_i^h\right).$$

Thus, the (i,j)-matrix entry is

$$(A)_{ij} = \int_\Omega \left(\boldsymbol{u}^{h,(m)} \cdot \nabla\right) \boldsymbol{\phi}_j^h \cdot \boldsymbol{\phi}_i^h \, d\boldsymbol{x} = \sum_{k=1}^{3} \int_\Omega \left(\left(\boldsymbol{u}^{h,(m)} \cdot \nabla\right) \boldsymbol{\phi}_j^h\right)_k \left(\boldsymbol{\phi}_i^h\right)_k \, d\boldsymbol{x}. \tag{6.83}$$

The product $\left(\left(\boldsymbol{u}^{h,(m)} \cdot \nabla\right) \boldsymbol{\phi}_j^h\right)_k \left(\boldsymbol{\phi}_i^h\right)_k$ vanishes if the kth component of $\boldsymbol{\phi}_j^h$ or the kth component of $\boldsymbol{\phi}_i^h$ vanishes. If $\boldsymbol{\phi}_i^h$ and $\boldsymbol{\phi}_j^h$ possess the same non-vanishing component, the matrix entry is independent of the component. Thus, the convective term in (6.82) leads to a matrix block of the velocity-velocity coupling of the form

$$A = \begin{pmatrix} A_{11} & 0 & 0 \\ 0 & A_{11} & 0 \\ 0 & 0 & A_{11} \end{pmatrix}. \tag{6.84}$$

- *Divergence form of the convective term.* In addition to matrix entries of form (6.83), the divergence form introduces the following entries to $(A)_{ij}$

$$\frac{1}{2} \int_\Omega \left(\nabla \cdot \boldsymbol{u}^{h,(m)}\right) \boldsymbol{\phi}_j^h \cdot \boldsymbol{\phi}_i^h \, d\boldsymbol{x} = \frac{1}{2} \sum_{k=1}^{3} \int_\Omega \left(\nabla \cdot \boldsymbol{u}^{h,(m)}\right) \left(\boldsymbol{\phi}_j^h\right)_k \left(\boldsymbol{\phi}_i^h\right)_k \, d\boldsymbol{x}.$$

A contribution of this type occurs also for the Galerkin discretization of the Oseen equations, see Remark 5.19. It is obvious that this entry is zero if the non-vanishing components of $\boldsymbol{\phi}_i$ and $\boldsymbol{\phi}_j$ are not the same. Otherwise, the value of this entry is independent of the index k of the non-vanishing entry. Altogether, the matrix of the velocity-velocity coupling has the form (6.84).

- *Rotational form of the convective term.* The matrix entry for the rotational form is given by

$$
\begin{aligned}
(A)_{ij} &= \left(\left(\nabla \times \boldsymbol{u}^{h,(m)}\right) \times \boldsymbol{\phi}_j^h, \boldsymbol{\phi}_i^h\right) \\
&= \int_\Omega \begin{pmatrix} \left(\nabla \times \boldsymbol{u}^{h,(m)}\right)_2 \left(\boldsymbol{\phi}_j^h\right)_3 - \left(\nabla \times \boldsymbol{u}^{h,(m)}\right)_3 \left(\boldsymbol{\phi}_j^h\right)_2 \\ \left(\nabla \times \boldsymbol{u}^{h,(m)}\right)_3 \left(\boldsymbol{\phi}_j^h\right)_1 - \left(\nabla \times \boldsymbol{u}^{h,(m)}\right)_1 \left(\boldsymbol{\phi}_j^h\right)_3 \\ \left(\nabla \times \boldsymbol{u}^{h,(m)}\right)_1 \left(\boldsymbol{\phi}_j^h\right)_2 - \left(\nabla \times \boldsymbol{u}^{h,(m)}\right)_2 \left(\boldsymbol{\phi}_j^h\right)_1 \end{pmatrix} \cdot \begin{pmatrix} \left(\boldsymbol{\phi}_i^h\right)_1 \\ \left(\boldsymbol{\phi}_i^h\right)_2 \\ \left(\boldsymbol{\phi}_i^h\right)_3 \end{pmatrix} d\boldsymbol{x}.
\end{aligned}
$$

  It follows that even if the non-vanishing components of $\boldsymbol{\phi}_i^h$ and $\boldsymbol{\phi}_j^h$ are different, the resulting entry will not vanish. Hence, the matrix for the velocity-velocity coupling has the form

$$
A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}, \tag{6.85}
$$

  which is the general form of this matrix.
- *Skew-symmetric form of the convective term.* Besides the half of the term (6.83), the skew-symmetric form possesses the contribution

$$
-\frac{1}{2}\int_\Omega \left(\boldsymbol{u}^{h,(m)} \cdot \nabla\right) \boldsymbol{\phi}_i^h \cdot \boldsymbol{\phi}_j^h \, d\boldsymbol{x} = -\frac{1}{2}\sum_{k=1}^{3} \int_\Omega \left(\left(\boldsymbol{u}^{h,(m)} \cdot \nabla\right) \boldsymbol{\phi}_i^h\right)_k \left(\boldsymbol{\phi}_j^h\right)_k \, d\boldsymbol{x}
$$

  in $(A)_{ij}$. From the same discussion as for (6.83), it follows that the matrix of the velocity-velocity coupling has the block-diagonal form (6.84).

Summary Using the convective form, the divergence form, or the skew-symmetric form of the convective term leads to block-diagonal matrices of form (6.84) in the Picard iteration. Only the rotational form requires the use of a full block matrix of form (6.85). □

Remark 6.43 (Newton's Method) In Newton's method, one takes as linear operator the derivative of the nonlinear operator at the current iterate

$$
\boldsymbol{N}_{\text{lin}} = D\boldsymbol{N}\begin{pmatrix} \boldsymbol{u}^{(m)} \\ p^{(m)} \end{pmatrix}.
$$

Considering the Gâteaux derivative, one obtains, using the linearity of $\boldsymbol{N}$ in each argument,

$$
\begin{aligned}
D\boldsymbol{N}\begin{pmatrix} \boldsymbol{u} \\ p \end{pmatrix} &= \lim_{\varepsilon \to 0} \frac{\boldsymbol{N}\left(\boldsymbol{u} + \varepsilon\boldsymbol{\phi}; \boldsymbol{u} + \varepsilon\boldsymbol{\phi}, p + \varepsilon\psi\right) - \boldsymbol{N}\left(\boldsymbol{u}; \boldsymbol{u}, p\right)}{\varepsilon} \\
&= \boldsymbol{N}\left(\boldsymbol{\phi}; \boldsymbol{u}, p\right) + \boldsymbol{N}\left(\boldsymbol{u}; \boldsymbol{\phi}, p\right) + \boldsymbol{N}\left(\boldsymbol{u}, \boldsymbol{u}, \psi\right).
\end{aligned}
$$

Using this operator as $\boldsymbol{N}_{\text{lin}}$ in (6.79) leads to

$$\begin{aligned}
&\boldsymbol{N}\Big(\tilde{\boldsymbol{u}}^{(m+1)};\boldsymbol{u}^{(m)},p^{(m)}\Big)+\boldsymbol{N}\Big(\boldsymbol{u}^{(m)};\tilde{\boldsymbol{u}}^{(m+1)},p^{(m)}\Big)+\boldsymbol{N}\Big(\boldsymbol{u}^{(m)},\boldsymbol{u}^{(m)},\tilde{p}^{(m+1)}\Big)\\
&\quad=\begin{pmatrix}\langle \boldsymbol{f},\boldsymbol{v}\rangle_{V',V}\\ 0\end{pmatrix}-\boldsymbol{N}\left(\boldsymbol{u}^{(m)};\boldsymbol{u}^{(m)},p^{(m)}\right)+\boldsymbol{N}\left(\boldsymbol{u}^{(m)};\boldsymbol{u}^{(m)},p^{(m)}\right)\\
&\qquad+\boldsymbol{N}\left(\boldsymbol{u}^{(m)};\boldsymbol{u}^{(m)},p^{(m)}\right)+\boldsymbol{N}\left(\boldsymbol{u}^{(m)},\boldsymbol{u}^{(m)},p^{(m)}\right).
\end{aligned}$$

Collecting terms gives

$$\begin{aligned}
&\begin{pmatrix}a\Big(\tilde{\boldsymbol{u}}^{(m+1)},\boldsymbol{v}\Big)+n\Big(\boldsymbol{u}^{(m)},\tilde{\boldsymbol{u}}^{(m+1)},\boldsymbol{v}\Big)+n\Big(\tilde{\boldsymbol{u}}^{(m+1)},\boldsymbol{u}^{(m)},\boldsymbol{v}\Big)+b\big(\boldsymbol{v},\tilde{p}^{(m+1)}\big)\\ b\left(\tilde{\boldsymbol{u}}^{(m+1)},q\right)\end{pmatrix}\\
&\qquad\qquad=\begin{pmatrix}\langle \boldsymbol{f},\boldsymbol{v}\rangle_{V',V}+n\left(\boldsymbol{u}^{(m)},\boldsymbol{u}^{(m)},\boldsymbol{v}\right)\\ 0\end{pmatrix}. \qquad (6.86)
\end{aligned}$$

In this method, the matrix and the right-hand side change in every iteration. Problem (6.86) is an Oseen problem with $\boldsymbol{b} = \boldsymbol{u}^{(m)}$ and the tensor-valued reaction $\nabla\boldsymbol{u}^{(m)}$. □

Remark 6.44 (On Newton's Method)

- The order of convergence of Newton's method is expected to be better than of the Picard iteration if the linear systems (6.86) are solved sufficiently accurately.
- Newton's method involves the reactive term $\left(\left(\boldsymbol{u}^{(m+1)}\cdot\nabla\right)\boldsymbol{u}^{(m)},\boldsymbol{v}\right)$ on the left-hand side. This term does not fit into the theory of the Oseen equations, since the required non-negativity of this term stated in Remark 5.2 is generally not given.

□

Remark 6.45 (Newton's Method: Structure of Matrices) The application of an inf-sup stable finite element method to (6.86) leads to a linear saddle point problem of form (6.81).

- *Convective form of the convective term.* The term

$$\begin{aligned}
&n\left(\boldsymbol{u}^{h,(m)},\boldsymbol{u}^{h,(m+1)},\boldsymbol{v}^h\right)+n\left(\boldsymbol{u}^{h,(m+1)},\boldsymbol{u}^{(h,m)},\boldsymbol{v}^h\right)\\
&\quad=\left(\left(\boldsymbol{u}^{h,(m)}\cdot\nabla\right)\boldsymbol{u}^{h,(m+1)},\boldsymbol{v}^h\right)+\left(\left(\boldsymbol{u}^{h,(m+1)}\cdot\nabla\right)\boldsymbol{u}^{h,(m)},\boldsymbol{v}^h\right)
\end{aligned}$$

arises on the left-hand side of the equation. The corresponding matrix entry for the convective form of the convective term becomes

$$(A)_{ij} = \sum_{k=1}^{3} \int_{\Omega} \left[\left(\boldsymbol{u}^{h,(m)} \cdot \nabla \left(\boldsymbol{\phi}_j^h\right)\right)_k \cdot \left(\boldsymbol{\phi}_i^h\right)_k \right.$$
$$\left. + \left(\sum_{l=1}^{3} \left(\boldsymbol{\phi}_j^h\right)_l \cdot \nabla \left(\boldsymbol{u}^{h,(m)}\right)_l\right)_k \cdot \left(\boldsymbol{\phi}_i^h\right)_k \right] d\boldsymbol{x}, \quad i,j = 1,\ldots,dN_v.$$

The sum in the second term is generally not zero for each index k since generally $\nabla\left(\boldsymbol{u}^{h,(m)}\right)_l$ does not vanish. It follows that the second term in general does not vanish for each pair of indices (i,j) with $\left|\operatorname{supp}\left(\boldsymbol{\phi}_i^h\right) \cap \operatorname{supp}\left(\boldsymbol{\phi}_j^h\right)\right| > 0$. Hence, the matrix blocks A_{kl} possess in general non-zero entries for all pairs (k,l). Altogether, the matrix which originates from the convective and reactive term in (6.86) has the block form (6.85), where the blocks are in general mutually different since different derivatives of $\boldsymbol{u}^{h,(m)}$ have to be considered in the assembling of each block.

- *Divergence form and skew-symmetric form of the convective term.* Both forms contain the term from the convective form. Since this term leads already to the matrix (6.85) of the velocity-velocity coupling, also the matrices for these two forms are of this type.
- *Rotational form of the convective term.* Already for the Picard iteration, the rotational form of the convective term requires the general form (6.85) of a matrix for the velocity-velocity coupling. With the additional term that is introduced from Newton's method, one gets the same form.

Summary The application of Newton's method leads always to a matrix for the velocity-velocity coupling of form (6.85) with mutually different blocks. □

Remark 6.46 (Memory Requirements and Computational Costs for the Velocity-Velocity Coupling) The matrix of the velocity-velocity coupling in the linear saddle point problems (6.80) and (6.86) is the sum of the matrix which arises in the discretization of the viscous term, see (4.93), and the matrix from the linearization of the convective term. Using the gradient form of the viscous term and the convective or divergence or skew-symmetric form of the convective term, then the velocity coupling is of form (6.84), such that only one matrix block needs to be stored. If Newton's method is employed, always the general form (6.85) of the velocity-velocity coupling is obtained. Note that with the appearance of more matrix blocks, the costs for assembling and for performing matrix-vector products increase. □

Example 6.47 (Picard Iteration vs. Newton's Method) In each step of the Picard iteration, a linear system of equations of form (6.80) has to be solved and in each step of Newton's method, a linear system of equations of form (6.86). There are two ways for performing this task:

- The linear system is solved exactly (up to round-off errors), e.g., with one of the sparse direct solvers mentioned in Remark 9.5.
- An iterative method is applied for the solution of the linear system. In this case, a usual approach consists in stopping the iteration after a few steps or after the reduction of the Euclidean norm of the residual vector by a prescribed factor. In this way, one performs an inexact solution of the linear system of equations.

For illustrating the behavior of the Picard iteration and of Newton's method with exact and inexact solutions of the arising linear systems of equations, the regularized driven cavity problem, Example D.4, with different values for the viscosity ν is considered. The simulations were performed for the Q_2/P_1^{disc} pair of finite element spaces on a grid consisting of squares (level 0 with four squares). The iteration on each level was started with a finite element velocity that took the value zero in all degrees of freedom in Ω. The inexact solution process stopped after having reduced the Euclidean norm of the residual vector by the factor 10 and at most 10 iterations were performed. The prescribed stopping criterion for the iteration required that the Euclidean norm of the residual vector should be smaller than 10^{-10}.

Results for the simulations are presented in Table 6.1. From the general theory of iterative schemes for nonlinear problems, it is known that the convergence radius of the Picard iteration (simple fixed point iteration) is generally larger than for

Table 6.1 Example 6.47

		Picard iteration		Newton's method	
ν	Level/lin. solver	Inexact	Exact	Inexact	Exact
1/100	2	14	26	8	5
	3	15	14	7	5
	4	14	14	7	5
	5	13	13	7	5
	6	13	13	7	5
1/500	2	39	Not conv.	Not conv.	Not conv.
	3	32	32	Not conv.	Not conv.
	4	30	29	35	8
	5	29	28	52	8
	6	28	27	Not conv.	8
1/1000	2	Not conv.	Not conv.	Not conv.	Not conv.
	3	36	57	Not conv.	Not conv.
	4	35	33	Not conv.	Not conv.
	5	35	31	Not conv.	Not conv.
	6	33	30	Not conv.	Not conv.

Number of iterations for solving the nonlinear problem, 'not conv.' means that the solution was not obtained within 100 iterations

Newton's method. Within the convergence radius, Newton's method converges faster (of second order) than the Picard method (of first order).

For a large value of the viscosity, $\nu = 100$, the faster convergence of Newton's method can be seen clearly. Using inexact solutions for the linear system of Eq. (6.86), Newton's method needs a few iterations more compared with using exact solutions.

For smaller values of ν, the convergence radius of Newton's method becomes obviously smaller. For $\nu = 1/500$, one can see that the convergence radius depends also on the fineness of the mesh. The considerably large number of iterations for the Picard method is because the stopping criterion that only 10 iterations should be performed became generally effective and the norm of the residual was reduced less than by a factor of 10. Obviously, it was more difficult for the iterative methods to solve the linear problems (6.86) of Newton's method than the linear problems (6.80) of the Picard method.

In the case $\nu = 1/1000$, it is probably not possible to resolve all important flow structures on level 2 such that all approaches did not converge. Apart from level 2, the Picard method worked for all considered values of the viscosity. The number of iterations was generally similar for the inexact and the exact solution of the linear systems of equations.

Altogether, it can be observed that in the case Newton's method converges, then the number of iterations is considerably smaller than for the Picard method. However, the converge radius might be small, in particular for small values of the viscosity. In this situation, the Picard iteration often still works. Of course, it is possible to construct iterative schemes which start with a method with larger convergence radius (Picard iteration, damped Newton's method) and which change to Newton's method once the iterates are sufficiently close to the solution. □

6.4 A Posteriori Error Estimation with the Dual Weighted Residual (DWR) Method

Remark 6.48 (Motivation) A posteriori error estimates, like the residual-based estimate (4.108) for the Stokes equations, possess two drawbacks. First, such error estimates can be derived usually only for norms that are in some sense natural in the setup of the problem, like the norm in V or the $L^2(\Omega)$ norm of the velocity or pressure. The error in such norms is generally not of much interest in applications. There, rather errors of drag and lift coefficients or other quantities of interest are important. And second, a posteriori error estimates of type (4.108) have still an unknown factor on the right-hand side, like C in (4.108). This factor contains in particular contributions from the stability of the problem and of local interpolation error estimates, compare Remark 4.80. The interpolation error depends of course on the finite element space, which in turn depends on the underlying grid. The stability of the problem usually depends on coefficients of the problem and this dependency might be severe, like for convection-dominated convection-diffusion equations, see John (2000) for numerical studies that reveal such dependencies. All these unknown dependencies might lead to constants that differ much from 1. In such cases, the knowledge of only the computable factor on the right-hand side of residual-based a posteriori error estimates is solely of limited use, since it does not allow to draw reliable conclusions on the actual size of the error.

The dual weighted residual (DWR) method is an approach which deals with both drawbacks at the same time. It is a widely applicable approach that leads to error estimates for functionals of interest. In a first step, an abstract representation of the error for a functional of interest is derived, see Lemma 6.51. Then, the abstract framework is applied to variational problems, which leads to an error representation that contains a primal residual that involves the solution of the discretized dual of a linearized problem, a dual residual that involves the finite element solution, and a remainder, see Theorem 6.54. For linear problems and linear functionals, it is shown in Remark 6.55 that the dual residual can be removed from the error representation. In other cases, the approximation of the dual residual might be computationally expensive since the argument is the solution of a nonlinear problem. Therefore, a second error representation is derived in Theorem 6.58 that contains only the primal residual and a quadratic remainder. In practice, the remainder is considered to be of higher order and the arguments for the evaluation of the residual will be approximated. Despite the approximations applied in this methodology, the experience is that the obtained estimates are usually close to the errors.

The DWR method was proposed in Becker and Rannacher (1996, 1998), reviews as well as references to previous papers considering a posteriori estimates for functionals of interest can be found in Becker and Rannacher (2001) and Bangerth and Rannacher (2003). □

Remark 6.49 (Abstract Setting) First, a general paradigm of a posteriori error analysis that is based on duality will be discussed. To this end, consider a differentiable functional $L(\cdot)$ defined on some linear space X. A stationary point of this functional is a point $x \in X$ for which

$$L'(x)(y) = 0 \quad \forall\, y \in X, \tag{6.87}$$

where the prime refers to the first argument. The Galerkin approximation of this problem reads as follows: Find $x^h \in X^h \subset X$ such that

$$L'\left(x^h\right)\left(y^h\right) = 0 \quad \forall\, y^h \in X^h. \tag{6.88}$$

The second argument of the derivative and of all higher order derivatives is linear. □

Example 6.50 (Energy Functional) Let X be a Hilbert space whose inner product is given by the bilinear form $a(\cdot,\cdot)$. Then the energy functional is defined by

$$L(x) = \frac{1}{2}a(x,x) - f(x) \quad x \in X,$$

where $f(\cdot)$ is a continuous, linear functional on X. Consider the function

$$\Phi(\varepsilon) = L(x + \varepsilon y) \quad \forall\, y \in X.$$

Then, $x \in X$ is a stationary point if

$$\Phi'(0) = 0 \quad \Longleftrightarrow \quad L'(x)(y) = 0 \quad \forall\, y \in X.$$

A straightforward calculation, using the linearity of $f(\cdot)$, gives

$$\begin{aligned}
\Phi'(0) &= \lim_{\varepsilon \to 0} \frac{\Phi(\varepsilon) - \Phi(0)}{\varepsilon} \\
&= \lim_{\varepsilon \to 0} \frac{\frac{1}{2}a(x,x) + \varepsilon a(x,y) - \frac{\varepsilon^2}{2}a(y,y) - f(x) - \varepsilon f(y) - \frac{1}{2}a(x,x) + f(x)}{\varepsilon} \\
&= \lim_{\varepsilon \to 0} \left(a(x,y) - f(y) - \frac{\varepsilon}{2}a(y,y) \right) \\
&= a(x,y) - f(y) \quad \forall\, y \in X.
\end{aligned}$$

Thus, a stationary point has to satisfy the equation

$$a(x,y) = f(y) \quad \forall\, y \in X \quad \Longleftrightarrow \quad L'(x)(y) = 0 \quad \forall\, y \in X.$$

Computing the second variation, one finds that the solution of the equation for a stationary point x is a minimizer of the energy functional. □

Lemma 6.51 (Abstract Error Representation) *Let the functional $L(\cdot)$ be three times differentiable. Then, for any solutions x of (6.87) and x^h of (6.88) there holds the error representation*

$$L(x) - L\left(x^h\right) = \frac{1}{2}L'\left(x^h\right)\left(x - I^h x\right) + \mathcal{R}, \tag{6.89}$$

where $I^h x \in X^h$ is arbitrary and with $e = x - x^h$ the remainder is given by

$$\mathcal{R} = \frac{1}{2}\int_0^1 \tau(\tau - 1)L'''\left(x^h + \tau e\right)(e,e,e)\,d\tau. \tag{6.90}$$

Proof With the fundamental theorem of calculus, one obtains

$$\int_0^1 L'\left(x^h+\tau e\right)(e)\,d\tau = L\left(x^h+\tau e\right)\Big|_{\tau=0}^{\tau=1} = L\left(x^h+e\right)-L\left(x^h\right)$$
$$= L(x)-L\left(x^h\right). \qquad (6.91)$$

In the next step, the integral will be represented with the help of the trapezoidal rule. To motivate this representation, let f be a sufficiently smooth function in $[t_0, t_1]$. Then, the truncation error in the trapezoidal rule is given by

$$\int_{t_0}^{t_1} f(\tau)\,d\tau - \frac{f(t_1)+f(t_0)}{2}(t_1-t_0).$$

Applying twice integration by parts and denoting $\bar{t} = (t_0+t_1)/2$ yields

$$\frac{1}{2}\int_{t_0}^{t_1}\left((\tau-\bar{t})^2-\left(\frac{t_1-t_0}{2}\right)^2\right)f''(\tau)\,d\tau = \int_{t_0}^{t_1}(\bar{t}-\tau)f'(\tau)\,d\tau$$
$$= \int_{t_0}^{t_1} f(\tau)\,d\tau - \frac{f(t_1)+f(t_0)}{2}(t_1-t_0).$$

Applying the trapezoidal rule to the left-hand side of (6.91) and using the derived form of the truncation error, i.e., setting $t_1 = 1$, $t_0 = 0$, gives

$$\int_0^1 L'\left(x^h+\tau e\right)(e)\,d\tau = \frac{L'(x)(e)+L'\left(x^h\right)(e)}{2}$$
$$+\frac{1}{2}\int_0^1 \tau(\tau-1)L'''\left(x^h+\tau e\right)(e,e,e)\,d\tau.$$

Since $e \in X$, one gets from (6.87) that $L'(x)(e) = 0$. Finally, the linearity of the second argument of the derivative and (6.88) leads for any $I^h x \in X^h$ to

$$L'\left(x^h\right)(e) = L'\left(x^h\right)\left(x-I^hx\right)+L'\left(x^h\right)\left(I^hx-x^h\right) = L'\left(x^h\right)\left(x-I^hx\right),$$

which completes the proof of the lemma. ■

Remark 6.52 (Interpretation of the Representation (6.89)*)* The identity (6.89) shows that the error for the functional can be represented with the operator and the solution of the Galerkin problem (6.88), and a remainder which is cubic in the error. Note that the operator of the Galerkin problem (6.88) is linear. However, the test function of the Galerkin operator contains the solution of the continuous problem (6.87), which is also a linear problem. □

Remark 6.53 (Application of the Basic Approach to a Variational Equation) Let V be a linear space and consider the equation: Find $u \in V$ such that

$$a(u; v) = a(u)(v) = f(v) \quad \forall\ v \in V, \tag{6.92}$$

where $a(\cdot)(\cdot)\ :\ V \times V \to \mathbb{R}$ is a differentiable form which is linear in the second argument and $f(\cdot)\ :\ V \to \mathbb{R}$ is a continuous linear functional. Let $V^h \subset V$ be a subspace, then the Galerkin approximation of (6.92) reads as follows: Find $u^h \in V^h$ such that

$$a\left(u^h\right)\left(v^h\right) = f\left(v^h\right) \quad \forall\ v^h \in V^h. \tag{6.93}$$

Denote by $J(\cdot)\ :\ V \to \mathbb{R}$ the functional of interest whose error $J(u) - J\left(u^h\right)$ should be minimized. To imbed this problem in the general framework derived so far, one considers the optimization problem

$$J(u) \to \min, \quad a(u)(v) = f(v) \quad \forall\ v \in V.$$

If (6.92) has a unique solution, then the optimization problem is trivial since in this case there is just one argument for the functional. This methodology can be also applied for problems with non-unique solutions, like eigenvalue problems.

Next, the Euler–Lagrange approach for deriving conditions for the solution of the optimization problem is applied. To this end, one considers the Lagrangian functional

$$\mathcal{L}(u, z) = J(u) + f(z) - a(u)(z),$$

where $z \in V$ is called adjoint variable. A necessary condition for a minimizer is that it is a stationary point, i.e., there holds

$$\begin{aligned} 0 &= \partial_u \mathcal{L}(u, z) = J'(u)(w) - a'(u)(w, z) \quad \forall\ w \in V, \\ 0 &= \partial_z \mathcal{L}(u, z) = f(v) - a(u)(v) \quad \forall\ v \in V. \end{aligned} \tag{6.94}$$

In the second relation, the linearity of $f(\cdot)$ was used such that the second condition is just (6.92). Equation (6.94) is called the dual problem associated to the functional $J(\cdot)$. Again, one considers the Galerkin approximations of the two conditions: Find $\left(u^h, z^h\right) \in V^h \times V^h$ such that

$$\begin{aligned} a'\left(u^h\right)\left(w^h, z^h\right) &= J'\left(u^h\right)\left(w^h\right) \ \forall\ w^h \in V^h, \\ a\left(u^h\right)\left(v^h\right) &= f\left(v^h\right) \qquad \forall\ v^h \in V^h. \end{aligned} \tag{6.95}$$

With the Galerkin solution, one defines the so-called primal residual

$$\rho\ :\ V \to V', \quad \rho(\cdot) = f(\cdot) - a\left(u^h\right)(\cdot), \tag{6.96}$$

and the dual residual

$$\rho^* \ : \ V \to V', \quad \rho^*(\cdot) = J'\left(u^h\right)(\cdot) - a'\left(u^h\right)\left(\cdot, z^h\right). \tag{6.97}$$

□

Theorem 6.54 (Error Representation) *Let the form $a(\cdot)(\cdot)$ and the functional $J(\cdot)$ be three times differentiable, let (u, z) be any solution of (6.94) and let $\left(u^h, z^h\right)$ be any solution of (6.95). Then there is the error representation*

$$J(u) - J\left(u^h\right) = \frac{1}{2}\rho\left(z - I^h z\right) + \frac{1}{2}\rho^*\left(u - I^h u\right) + \mathcal{R}_a, \tag{6.98}$$

where $I^h z, I^h u \in V^h$ are arbitrary functions. Denoting $e = u - u^h$ and $e^* = z - z^h$, the remainder is given by

$$\mathcal{R}_a = \frac{1}{2}\int_0^1 \tau(\tau-1)\Big[J'''\left(u^h + \tau e\right)(e,e,e) - a'''\left(u^h + \tau e\right)\left(e,e,e,z^h + \tau e^*\right)$$
$$-3a''\left(u^h + \tau e\right)\left(e,e,e^*\right)\Big]\, d\tau.$$

Proof This situation will be imbedded in the general framework of Lemma 6.51. To this end, one sets $X = V \times V$, $X^h = V^h \times V^h$, $x = (u, z)$, $x^h = \left(u^h, z^h\right)$, and $L(x) = \mathcal{L}(u, z)$. Then, using the definition of the Lagrangian functional, the total derivative at x^h is given by

$$\begin{aligned} L'\left(x^h\right)(v_u, v_z) &= \partial_u \mathcal{L}\left(u^h, z^h\right)(v_u) + \partial_z \mathcal{L}\left(u^h, z^h\right)(v_z) \\ &= J'\left(u^h\right)(v_u) - a'\left(u^h\right)\left(v_u, z^h\right) + f(v_z) - a\left(u^h\right)(v_z) \end{aligned} \tag{6.99}$$

for arbitrary $(v_u, v_z) \in X$. Again, the linearity of $f(\cdot)$ and the second argument of $a(\cdot)(\cdot)$ was used. The application of (6.94), (6.95), (6.89), (6.99), (6.96), and (6.97) yields

$$\begin{aligned} &J(u) - J\left(u^h\right) \\ &= J(u) + f(z) - a(u)(z) - \left(J\left(u^h\right) + f\left(z^h\right) - a\left(u^h\right)\left(z^h\right)\right) \\ &= L(x) - L\left(x^h\right) \\ &= \frac{1}{2}L'\left(x^h\right)\left(x - I^h x\right) + \mathcal{R} \\ &= \frac{1}{2}\Big(J'\left(u^h\right)\left(u - I^h u\right) - a'\left(u^h\right)\left(u - I^h u, z^h\right) + f\left(z - I^h z\right) - a\left(u^h\right)\left(z - I^h z\right)\Big) + \mathcal{R} \\ &= \frac{1}{2}\Big(\rho\left(z - I^h z\right) + \rho^*\left(u - I^h u\right)\Big) + \mathcal{R}, \end{aligned}$$

with arbitrary $\left(I^h u, I^h z\right) \in X^h$. The general form (6.90) of the remainder requires to compute the third derivative of $L(x) = \mathcal{L}(u, z)$. Formally, one has

$$L''' = \partial_{uuu}\mathcal{L} + 3\partial_{uuz}\mathcal{L} + \partial_{uzz}\mathcal{L} + \partial_{zzz}\mathcal{L}.$$

Since $\mathcal{L}$ depends only linearly on z, the last two terms vanish. The derivative of the other two terms at the considered point $\left(x^h + \tau\left(e, e^*\right)\right)\left(\left(e, e^*\right), \left(e, e^*\right), \left(e, e^*\right)\right)$ gives just the terms in $\mathcal{R}_a$

$$\begin{aligned} &L'''\left(x^h + \tau\left(e, e^*\right)\right)\left(\left(e, e^*\right), \left(e, e^*\right), \left(e, e^*\right)\right) \\ &\quad = J'''\left(u^h + \tau e\right)(e, e, e) - a'''\left(u^h + \tau e\right)\left(e, e, e, z^h + \tau e^*\right) \\ &\qquad - 3a''\left(u^h + \tau e\right)\left(e, e, e^*\right). \end{aligned} \tag{6.100}$$

For the second term on the left-hand side, the three derivatives with respect to u lead to three times the argument e and for the last term, the two derivatives with respect to u and the last derivative with respect to z lead to twice the argument e and once the argument e^*. Note that the prime refers only to the first argument of $a(\cdot)(\cdot)$, which gives the last term in (6.100). ■

Remark 6.55 (Linear Variational Problem and Linear Functional) Consider a linear variational problem

$$a(u)(v) = a(u, v) = f(v) \quad \forall\ v \in V$$

and the corresponding Galerkin approximation

$$a\left(u^h, v^h\right) = f\left(v^h\right) \quad \forall\ v^h \in V^h.$$

Subtracting both equations, one gets the Galerkin orthogonality

$$a\left(u - u^h, v^h\right) = 0 \quad \forall\ v^h \in V^h.$$

Since $a'\left(u^h\right)\left(\cdot, z^h\right) = a\left(\cdot, z^h\right)$ and $J'\left(u^h\right)(\cdot) = J(\cdot)$, one obtains with the Galerkin orthogonality, since $z^h \in V^h$, the linearity of the functional, and $I^h u = u^h$

$$\rho^*\left(u - I^h u\right) = J(u) - a\left(u, z^h\right) - J\left(u^h\right) + a\left(u^h, z^h\right) = J\left(u - u^h\right) = J(e).$$

Inserting this expression in the error representation (6.98) and observing that the remainder vanishes, since all higher order derivatives of the variational form and the functional vanish, yields

$$J(u) - J\left(u^h\right) = J(e) = \rho\left(z - I^h z\right).$$

Thus, for this special case one needs to compute only the primal residual, evaluated for the difference of the solution of the dual problem and an arbitrary interpolation. □

Remark 6.56 (Motivation for Removing the Dual Residual from the Error Representation) Estimating the error $J(u) - J\left(u^h\right)$ for a nonlinear problem or a nonlinear functional with (6.98), by neglecting $\mathcal{R}_a$, requires the approximation of the primal and the dual residual. Note that the solutions u and z of the continuous problems are usually not available and they need to be approximated. These approximations must not be contained in V^h since otherwise the residuals vanish. A linear problem has to be solved for approximating z whereas the approximation of u requires the solution of a nonlinear problem. Thus, in particular the approximation of the dual residual might be computationally expensive and there is the desire to remove the dual residual from the error representation. It will be shown now that it is possible to represent the dual residual in terms of the primal residual plus a quadratic remainder term. □

Lemma 6.57 (Representation of the Dual Residual with the Primal Residual and a Quadratic Remainder) *Let the form $a(\cdot)(\cdot)$ and the functional $J(\cdot)$ be twice differentiable, let (u, z) be any solution of (6.94), and let $\left(u^h, z^h\right)$ be any solution of (6.95). Then it holds for arbitrary functions $I^h z, I^h u \in V^h$ that*

$$\rho^* \left(u - I^h u\right) = \rho \left(z - I^h z\right) + \delta\rho, \tag{6.101}$$

with

$$\delta\rho = \int_0^1 \left(a'' \left(u^h + \tau e\right) \left(e, e, z^h + \tau e^*\right) - J'' \left(u^h + \tau e\right) (e, e)\right) d\tau$$

and $e = u - u^h$, $e^* = z - z^h$.

Proof Defining the function

$$g(\tau) = J' \left(u^h + \tau e\right) (e) - a' \left(u^h + \tau e\right) \left(e, z^h + \tau e^*\right),$$

one finds with (6.94) that

$$g(1) = J' \left(u^h + e\right) (e) - a' \left(u^h + e\right) \left(e, z^h + e^*\right) = J' (u) (e) - a' (u) (e, z) = 0 \tag{6.102}$$

and with (6.97) that

$$g(0) = J' \left(u^h\right) (e) - a' \left(u^h\right) \left(e, z^h\right) = \rho^*(e). \tag{6.103}$$

For computing the derivative of $g(\tau)$, one uses the definition of the second derivative of J, the linearity of the second argument of the form a, and the definition of the second derivative of a

$$
\begin{aligned}
g'(\tau) &= \lim_{\varepsilon\to 0} \frac{g(\tau+\varepsilon)-g(\tau)}{\varepsilon} \\
&= \lim_{\varepsilon\to 0} \frac{J'\left(u^h+(\tau+\varepsilon)e\right)(e) - J'\left(u^h+\tau e\right)(e)}{\varepsilon} \\
&\quad - \lim_{\varepsilon\to 0} \frac{a'\left(u^h+(\tau+\varepsilon)e\right)\left(e, z^h+(\tau+\varepsilon)e^*\right) - a'\left(u^h+\tau e\right)\left(e, z^h+\tau e^*\right)}{\varepsilon} \\
&= J''\left(u^h+\tau e\right)(e,e) - \lim_{\varepsilon\to 0} \frac{\varepsilon a'\left(u^h+(\tau+\varepsilon)e\right)(e,e^*)}{\varepsilon} \\
&\quad - \lim_{\varepsilon\to 0} \frac{a'\left(u^h+(\tau+\varepsilon)e\right)\left(e, z^h+\tau e^*\right) - a'\left(u^h+\tau e\right)\left(e, z^h+\tau e^*\right)}{\varepsilon} \\
&= J''\left(u^h+\tau e\right)(e,e) - a'\left(u^h+\tau e\right)\left(e,e^*\right) - a''\left(u^h+\tau e\right)\left(e,e,z^h+\tau e^*\right).
\end{aligned}
$$

Applying the fundamental theorem of calculus, (6.102), and (6.103) gives

$$
\begin{aligned}
\rho^*(e) = g(0) - g(1) &= \int_1^0 g'(\tau)\,d\tau \\
&= \int_0^1 -J''\left(u^h+\tau e\right)(e,e)\,d\tau + \int_0^1 a''\left(u^h+\tau e\right)\left(e,e,z^h+\tau e^*\right)\,d\tau \\
&\quad + \int_0^1 a'\left(u^h+\tau e\right)\left(e,e^*\right)\,d\tau. \qquad (6.104)
\end{aligned}
$$

For the last term, one obtains with the definition (6.92) of the form a and the definition of the primal residual (6.96)

$$
\begin{aligned}
\int_0^1 a'\left(u^h+\tau e\right)\left(e,e^*\right)\,d\tau &= a\left(u^h+\tau e\right)\left(e^*\right)\Big|_{\tau=0}^{\tau=1} = a\left(u^h+e\right)\left(e^*\right) - a\left(u^h\right)\left(e^*\right) \\
&= a(u)\left(e^*\right) - a\left(u^h\right)\left(e^*\right) = f\left(e^*\right) - a\left(u^h\right)\left(e^*\right) \\
&= \rho\left(e^*\right) = \rho\left(z-z^h\right). \qquad (6.105)
\end{aligned}
$$

Using the linearity of f and of the form a with respect to the second argument and the definition (6.93) of the Galerkin approximation yields

$$
\begin{aligned}
\rho\left(z-z^h\right) &= f\left(z-z^h\right) - a\left(u^h\right)\left(z-z^h\right) = f(z) - a\left(u^h\right)(z) - \left(f\left(z^h\right) - a\left(u^h\right)\left(z^h\right)\right) \\
&= f(z) - a\left(u^h\right)(z) - \left(f\left(I^h z\right) - a\left(u^h\right)\left(I^h z\right)\right) = \rho\left(z - I^h z\right) \qquad (6.106)
\end{aligned}
$$

for arbitrary $I^h z \in V^h$. With the linearity of J' and a' with respect to the second argument and the Galerkin formulation of the conditions for minima (6.95), one gets in the same way

$$\rho^*(e) = \rho^*\left(u - u^h\right) = \rho^*\left(u - I^h u\right) \quad \forall\, I^h u \in V^h. \tag{6.107}$$

Inserting now (6.105), (6.106), and (6.107) in (6.104) gives the statement of the lemma. ■

Theorem 6.58 (Error Representation with the Primal Residual and Quadratic Remainder) *Let the assumptions of Theorem 6.54 be valid, then one has the error representation*

$$J(u) - J\left(u^h\right) = \rho\left(z - I^h z\right) + \widetilde{\mathcal{R}}_a \tag{6.108}$$

with $e = u - u^h$, $e^* = z - z^h$, and

$$\widetilde{\mathcal{R}}_a = \int_0^1 \tau\Big(a''\left(u^h + \tau e\right)(e,e,z) - J''\left(u^h + \tau e\right)(e,e)\Big)\, d\tau.$$

Proof Inserting (6.101) in (6.98) gives with a straightforward calculation

$$J(u) - J\left(u^h\right) = \rho\left(z - I^h z\right) + \mathcal{R}_a + \frac{1}{2}\delta\rho.$$

Thus, one has to show that $\widetilde{\mathcal{R}}_a - \mathcal{R}_a - (\delta\rho)/2 = 0$.

Integration by parts, the linearity of the second argument of the form a, and the product rule gives

$$\begin{aligned}
\frac{1}{2}\delta\rho = {} & \left[\frac{\tau}{2}\Big(a''\left(u^h + \tau e\right)\left(e, e, z^h + \tau e^*\right) - J''\left(u^h + \tau e\right)(e,e)\Big)\right]_{\tau=0}^{\tau=1} \\
& -\frac{1}{2}\int_0^1 \tau\Big[a'''\left(u^h + \tau e\right)\left(e, e, e, z^h + \tau e^*\right) + a''\left(u^h + \tau e\right)\left(e, e, e^*\right) \\
& -J'''\left(u^h + \tau e\right)(e,e,e)\Big]\, d\tau \\
= {} & \frac{1}{2}\left(a''(u)(e,e,z) - J''(u)(e,e)\right) - \frac{1}{2}\int_0^1 \tau\Big[a'''\left(u^h + \tau e\right)\left(e,e,e,z^h + \tau e^*\right) \\
& + a''\left(u^h + \tau e\right)\left(e,e,e^*\right) - J'''\left(u^h + \tau e\right)(e,e,e)\Big]\, d\tau.
\end{aligned} \tag{6.109}$$

Likewise, one obtains with integration by parts

$$\begin{aligned}
\widetilde{\mathcal{R}}_a = {} & \left[\frac{\tau^2}{2}\Big(a''\left(u^h + \tau e\right)(e,e,z) - J''\left(u^h + \tau e\right)(e,e)\Big)\right]_{\tau=0}^{\tau=1} \\
& - \int_0^1 \frac{\tau^2}{2}\Big(a'''\left(u^h + \tau e\right)(e,e,e,z) - J'''\left(u^h + \tau e\right)(e,e,e)\Big)\, d\tau
\end{aligned}$$

$$= \frac{1}{2}\left(a''(u)(e,e,z) - J''(z)(e,e)\right) - \int_0^1 \frac{\tau^2}{2}\Big(a'''\left(u^h + \tau e\right)(e,e,e,z) - J'''\left(u^h + \tau e\right)(e,e,e)\Big)\,d\tau. \tag{6.110}$$

The terms outside the integrals in (6.109) and (6.110) cancel. It remains to show that the terms in the integrals of $\widetilde{\mathcal{R}}_a - \mathcal{R}_a - (\delta\rho)/2$ vanish. For the part with the third derivative of the functional, it is obvious since

$$\frac{\tau^2}{2}J'''\left(u^h + \tau e\right)(e,e,e) + \frac{\tau(\tau-1)}{2}J'''\left(u^h + \tau e\right)(e,e,e) + \frac{\tau}{2}J'''\left(u^h + \tau e\right)(e,e,e) = 0.$$

For the form a, one obtains with (6.109), (6.110), $-z + z^h + \tau e^* = -e^* + \tau e^*$, integration by parts, and the linearity of the a with respect to the second argument

$$\begin{aligned}
&\int_0^1 -\frac{\tau^2}{2}a'''\left(u^h + \tau e\right)(e,e,e,z) + \frac{\tau(\tau-1)}{2}\Big(a'''\left(u^h + \tau e\right)\left(e,e,e,z^h + \tau e^*\right) \\
&\quad +3a''\left(u^h + \tau e\right)\left(e,e,e^*\right)\Big) + \frac{\tau}{2}\Big(a'''\left(u^h + \tau e\right)\left(e,e,e,z^h + \tau e^*\right) \\
&\quad +a''\left(u^h + \tau e\right)\left(e,e,e^*\right)\Big)\,d\tau \\
&= \int_0^1 \frac{\tau^2}{2}a'''\left(u^h + \tau e\right)\left(e,e,e,-e^* + \tau e^*\right) + \frac{3\tau^2}{2}a''\left(u^h + \tau e\right)\left(e,e,e^*\right) \\
&\quad -\tau a''\left(u^h + \tau e\right)\left(e,e,e^*\right)\,d\tau \\
&= \int_0^1 \frac{\tau^2}{2}a'''\left(u^h + \tau e\right)\left(e,e,e,-e^* + \tau e^*\right) \\
&\quad +\frac{3}{2}\frac{\tau^3}{3}a''\left(u^h + \tau e\right)\left(e,e,e^*\right)\Big|_{\tau=0}^{\tau=1} - \int_0^1 \frac{3}{2}\frac{\tau^3}{3}a'''\left(u^h + \tau e\right)\left(e,e,e,e^*\right)\,d\tau \\
&\quad -\frac{\tau^2}{2}a''\left(u^h + \tau e\right)\left(e,e,e^*\right)\Big|_{\tau=0}^{\tau=1} + \int_0^1 \frac{\tau^2}{2}a'''\left(u^h + \tau e\right)\left(e,e,e,e^*\right)\,d\tau \\
&= \int_0^1 \Big[\frac{\tau^2}{2}a'''\left(u^h + \tau e\right)\left(e,e,e,-e^* + \tau e^*\right) - \frac{\tau^2}{2}a'''\left(u^h + \tau e\right)\left(e,e,e,\tau e^*\right) \\
&\quad + \frac{\tau^2}{2}a'''\left(u^h + \tau e\right)\left(e,e,e,e^*\right)\Big]\,d\tau + \frac{1}{2}a''(u)\left(e,e,e^*\right) - \frac{1}{2}a''(u)\left(e,e,e^*\right) \\
&= 0.
\end{aligned}$$

This identity finishes the proof of the theorem. ■

Remark 6.59 (Some Practical Aspects)

- In practice, the error is approximated with the representation (6.108)

$$J(u) - J\left(u^h\right) \approx \rho\left(z - I^h z\right), \tag{6.111}$$

  thus neglecting the quadratic remainder term. As shown in Remark 6.55, this approximation is for linear variational problems and linear functionals the same that is obtained from (6.98). In this case, a cubic remainder is neglected.
- The approximation of the right-hand side of (6.111) requires an approximation of the solution z of the dual linearized problem. As already discussed in Remark 6.56, this approximation must not be contained in V^h since in this case the residual vanishes. There are several proposals on how to computed this approximation, e.g., by using a higher order method or by post-processing a discrete approximation in V^h in a higher order finite element space, see Becker and Rannacher (2001, Sect. 5) and Bangerth and Rannacher (2003, Sect. 4.1) for details.
- The dual linearized problem for the steady-state Navier–Stokes equations is given in (6.72), see also (6.71), usually with a non-homogeneous right-hand side in the second equation, depending on the functional of interest. In contrast to the Navier–Stokes equations, the convection possesses the opposite sign, i.e., the main flow is in the opposite direction. This issue might lead to problems if an adaptive grid refinement is used for the Navier–Stokes equations, since a properly adapted grid for the Navier–Stokes equations will usually not be properly adapted for approximating the solution of the dual linearized problem.
- Typical quantities of interest for incompressible flow problems are the drag and lift coefficients of bodies in flow fields, see Remark D.2. Examples of adaptively refined grids with respect to these quantities of interest can be found, e.g., in Becker (2000) and (Becker and Rannacher 2001, Sect. 8).
- Another difficulty shows up for time-dependent problems. Applying the strategy for deriving the dual problem explained in Remark 4.27 to the term $(\partial_t \boldsymbol{u}, \boldsymbol{v})$ gives, using integration by parts in time, for the dual problem $-(\partial_t \boldsymbol{z}, \boldsymbol{v})$. Thus, the dual linearized problem is backward in time. Its simulation requires a computation starting from the current discrete time back to the initial time. The backward-in-time problem needs the computed velocity field at all former discrete times, since these fields are the input data $\boldsymbol{w}$ of the dual linearized problem (6.72). The storage of these velocity fields is memory-consuming. There are so-called checkpoint techniques that only store some of the velocity fields and that reconstruct the intermediate velocity fields between the checkpoints during the solution of the backward-in-time problem, e.g., see Besier and Rannacher (2012) and the references therein. □

Remark 6.60 (Classical Residual-Based A Posteriori Error Estimators) Residual-based a posteriori error estimators in the classical sense, compare Sect. 4.4 for the Stokes equations, can be derived also for the steady-state Navier–Stokes equations, e.g., see Verfürth (2013, Sect. 5.4.3). □

Chapter 7
The Time-Dependent Navier–Stokes Equations: Laminar Flows

Remark 7.1 (The Time-Dependent Navier–Stokes Equations) The time-dependent Navier–Stokes equations (2.25) were derived in Chap. 2 as a model for describing the behavior of incompressible fluids. From the point of view of numerical simulations, one has to distinguish between laminar and turbulent flows. It does not exist an exact definition of these terms. From the point of view of simulations, a flow is considered to be laminar, if on reasonable grids all flow structures can be represented or resolved. In this case, it is possible to simulate the flow with standard discretization techniques in space, like the Galerkin finite element method.

In addition to the discretization in space, the simulation of time-dependent flows requires a discretization of the temporal derivative of the velocity. There are a number of so-called time stepping schemes available that lead to a variety of different algorithms for the numerical simulation of the time-dependent Navier–Stokes equations. □

7.1 The Continuous Equations

Remark 7.2 (The Incompressible Navier–Stokes Equations) This chapter considers the incompressible Navier–Stokes equations given by

$$\begin{aligned} \partial_t \boldsymbol{u} - \nu \Delta \boldsymbol{u} + (\boldsymbol{u} \cdot \nabla)\boldsymbol{u} + \nabla p &= \boldsymbol{f} \quad \text{in } (0,T] \times \Omega, \\ \nabla \cdot \boldsymbol{u} &= 0 \quad \text{in } (0,T] \times \Omega, \\ \boldsymbol{u}(0,\cdot) &= \boldsymbol{u}_0 \text{ in } \Omega, \end{aligned} \tag{7.1}$$

where $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, is a domain. Only the case will be studied that Ω is a bounded domain with Lipschitz boundary Γ. In addition, Ω does not change in

V. John, *Finite Element Methods for Incompressible Flow Problems*, Springer Series in Computational Mathematics 51, DOI 10.1007/978-3-319-45750-5_7

time. For simplicity of presentation, homogeneous Dirichlet boundary conditions

$$\boldsymbol{u} = \boldsymbol{0} \text{ in } (0, T] \times \Gamma \tag{7.2}$$

are assumed. □

Remark 7.3 (Notions of a Solution) In the literature, several notions of a (velocity) solution of (7.1) that possess different properties (regularity) can be found, like classical, strong, mild, weak, or very weak solution, e.g., see Farwig (2014). This chapter studies the weak solution. □

Remark 7.4 (On the Weak Equation for the Velocity) In the case of the time-dependent incompressible Navier–Stokes equations, one restricts the analysis in the first step to an appropriate divergence-free subspace and studies the existence and uniqueness of an appropriate velocity. In the second step, the existence of a corresponding pressure is studied. Different forms of a weak formulation for the velocity can be found in the literature. □

Remark 7.5 (First Form of a Weak Equation for the Velocity) One considers, as for the stationary equations, test functions from V_{div}, multiplies the momentum equation in (7.1) with these test functions, applies integration by parts in space and obtains an ordinary differential equation for the velocity

$$\begin{aligned} \frac{d}{dt}(\boldsymbol{u}, \boldsymbol{v}) + (\nu \nabla \boldsymbol{u}, \nabla \boldsymbol{v}) + n(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}) &= \langle \boldsymbol{f}, \boldsymbol{v} \rangle_{V', V} \ \forall \ \boldsymbol{v} \in V_{\text{div}}, \\ \boldsymbol{u}(0) &= \boldsymbol{u}_0. \end{aligned} \tag{7.3}$$

Here, the derivative with respect to time has to be understood in a weak sense. This form of the weak equation can be found, e.g., in Girault and Raviart (1979, p. 158) and Temam (1984, p. 280). □

Definition 7.6 (Weak or Variational Velocity Solution of the Navier–Stokes Equations (Girault and Raviart 1979, p. 158; Temam 1984, p. 280)) Let $\boldsymbol{f} \in L^2(0, T; V')$ and $\boldsymbol{u}_0 \in H_{\text{div}}(\Omega)$. Then $\boldsymbol{u} \in L^2(0, T; V_{\text{div}})$ is called weak or variational velocity solution of the Navier–Stokes equations if $\boldsymbol{u}$ satisfies (7.3) in $(C_0^\infty((0, T)))'$. □

Remark 7.7 (Second Form of a Weak Equation for the Velocity) To obtain this form, integration with respect to time and integration by parts are applied additionally. One can consider this approach also in the way that the momentum equation of (7.1) is multiplied with test functions depending on time and space, it is integrated in the time-space domain, and then integration by parts is applied.

Usually, smooth test functions are used for this purpose. The extension of the statements to test function in appropriate Lebesgue and Sobolev spaces is then based on the density of the smooth functions in Lebesgue and Sobolev spaces, like stated in Theorem A.38. To be concrete, test functions from the space $C_{0,\text{div}}^\infty([0, T) \times \Omega)$ are applied. Since these functions are in $C_0^\infty([0, T))$ with respect to time, they vanish at the final time. Applying now the approach for deriving a weak formulation yields

$$\int_0^T \Big[-(\boldsymbol{u}, \partial_t \boldsymbol{\phi}) + \nu(\nabla \boldsymbol{u}, \nabla \boldsymbol{\phi}) + ((\boldsymbol{u}\cdot\nabla)\boldsymbol{u}, \boldsymbol{\phi}) \Big](\tau)\, d\tau \tag{7.4}$$

$$= \int_0^T \langle \boldsymbol{f}, \boldsymbol{\phi}\rangle_{V',V}(\tau)\, d\tau + (\boldsymbol{u}_0, \boldsymbol{\phi}(0,\cdot)) \quad \forall\, \boldsymbol{\phi} \in C^\infty_{0,\mathrm{div}}([0,T)\times\Omega).$$

Note that the temporal derivative is applied to the test function and not to the velocity. There is no contribution from the integration by parts in time at the final time because the test functions vanish at this time. Equation (7.4) is an integral form of the ordinary differential equation (7.3). A weak formulation of form (7.4) can be found, e.g., in Galdi (2000), Amann (2000), and Sohr (2001, p. 263). □

Definition 7.8 (Weak or Variational Solution of the Navier–Stokes Equations (Galdi 2000; Sohr 2001, p. 263)) Let $\boldsymbol{f} \in L^2(0,T;V')$ and $\boldsymbol{u}_0 \in H_{\mathrm{div}}(\Omega)$. A function $\boldsymbol{u}$ is called weak or variational solution of the Navier–Stokes equations if

- $\boldsymbol{u}$ satisfies (7.4),
- $\boldsymbol{u}$ has the following regularity

$$\boldsymbol{u} \in L^2(0,T;V_{\mathrm{div}}) \cap L^\infty(0,T;H_{\mathrm{div}}(\Omega)). \tag{7.5}$$

□

Remark 7.9 (General Idea for Proving Existence of a Weak Solution) The general idea to prove the existence of a weak solution is as follows:

1. Consider a sequence of simpler problems than (7.4) which in an appropriate limit converges to (7.4).
2. Show that each of the simpler problems has a unique solution.
3. Show that a subsequence of the sequence of the unique solutions converges to a weak solution of the Navier–Stokes equations.

□

Remark 7.10 (On the Definition of Simpler Problems) There are different proposals for defining simpler problems in the literature. Some of them are the followings.

- *Leray's regularization approach.* The first results using this approach were obtained by Leray:
  - 1933: $\Omega = \mathbb{R}^2$, Leray (1933),
  - 1934: Ω is a fixed oval in $\mathbb{R}^2$, Leray (1934a),
  - 1934: $\Omega = \mathbb{R}^3$, Leray (1934b).

  The most remarkable one is the last paper. There, the simplified equations have the convective term

$$(\boldsymbol{u}_\varepsilon \cdot \nabla)\,\boldsymbol{u},$$

instead of $(\boldsymbol{u}\cdot\nabla)\,\boldsymbol{u}$ in the Navier–Stokes equations (7.1), with $\boldsymbol{u}_\varepsilon$ being an average of $\boldsymbol{u}$ in space with averaging radius ε

$$\boldsymbol{u}_\varepsilon(\boldsymbol{x}) := \frac{1}{\varepsilon^3}\int_\Omega \lambda\left(\frac{\|\boldsymbol{x}-\boldsymbol{y}\|_2^2}{\varepsilon^2}\right)\boldsymbol{u}(\boldsymbol{y})\,d\boldsymbol{y}, \quad \lambda(s) = \lambda_0 e^{\frac{1}{s-1}},\ s \in (0,1),$$

where λ_0 is a constant for achieving an appropriate normalization. The use of the convection field $\boldsymbol{u}_\varepsilon$ corresponds to a regularization of the Navier–Stokes equations. In this approach, the limit $\varepsilon \to 0$ is studied.

- *Semidiscretization in space, Galerkin method.* In Hopf (1951), a different type of simpler problems to be considered in the first step of the general approach was introduced. In this method, (7.3) or (7.4) is considered in finite-dimensional spaces with dimension n, which is the so-called Galerkin method. That means, the equation has the same form as (7.3) or (7.4) but the test functions are from a finite-dimensional space and the solution is sought in the same space. Then, $n \to \infty$ is studied. The weak solution of the Navier–Stokes equations obtained with this approach is called weak solution in the sense of Leray–Hopf.
- *Semidiscretization in time.* It is also possible to define the simpler problems with a discretization in time, see Temam (1984, Chap. III.4). In the limit, one passes from the discrete times to the continuous time.
- *The semigroup method.* Another approach for proving the existence of a weak solution was developed in Sohr (1983), see also Sohr (2001). It is based on a semigroup, see Remark 8.16. Similarly to Leray's approach, the simpler problems rely on a regularization of the convective term. This regularization is performed with the Yosida approximation, which is the continuous counterpart of the differential filter used in turbulence modeling.
- *Turbulence models.* Turbulence models possess a smaller complexity than the Navier–Stokes equations to facilitate flow simulations, i.e., they define simpler problems. The limit of a turbulence model might also provide a weak solution of the Navier–Stokes equations, e.g., see Remark 8.203 for the Navier–Stokes-α model.

Below, an approach using the Galerkin method will be presented. □

Remark 7.11 (Starting Point of the Galerkin Method: The Navier–Stokes Equations in a Finite-Dimensional Subspace) The first step of the Galerkin method consists in considering the weak form of the Navier–Stokes equations (7.3) or (7.4) in a finite-dimensional space. It can be shown, see Galdi (2000, Lemma 2.3), that there is a basis of $\{\boldsymbol{v}_l\}_{l=1}^\infty$ of $C_{0,\mathrm{div}}^\infty(\Omega)$ where the basis functions are orthonormal with respect to the inner product of $L^2(\Omega)$. Consider now the finite-dimensional subspace

$$V_{\mathrm{div}}^n = \mathrm{span}\{\boldsymbol{v}_l^n\}_{l=1}^n \subset C_{0,\mathrm{div}}^\infty(\Omega).$$

In this subspace, the Galerkin method applied to (7.3) reads as follows: Find $\boldsymbol{u}^n \in V_{\text{div}}^n$ such that

$$(\partial_t \boldsymbol{u}^n, \boldsymbol{v}^n) + (\nu \nabla \boldsymbol{u}^n, \nabla \boldsymbol{v}^n) + n\,(\boldsymbol{u}^n, \boldsymbol{u}^n, \boldsymbol{v}^n) = \langle \boldsymbol{f}, \boldsymbol{v}^n \rangle_{V',V} \quad \forall\; \boldsymbol{v}^n \in V_{\text{div}}^n, \tag{7.6}$$

and $\boldsymbol{u}^n(0) = \boldsymbol{u}_0^n$, where $\boldsymbol{u}_0^n$ is the $L^2(\Omega)$ orthogonal projection of $\boldsymbol{u}_0$ into V_{div}^n. It is clear that (7.6) is satisfied if this equation is satisfied for all basis functions. Using the representation

$$\boldsymbol{u}^n(t, \boldsymbol{x}) = \sum_{l=1}^{n} \alpha_l^n(t) \boldsymbol{v}_l^n(\boldsymbol{x}), \tag{7.7}$$

one obtains from (7.6) the following system of ordinary differential equations

$$\frac{d\alpha_l^n}{dt} + \sum_{j=1}^{n} a_{lj}\alpha_j^n + \sum_{j,k=1}^{n} n_{ljk}\alpha_j^n\alpha_k^n = f_l, \quad l = 1, \ldots, n, \tag{7.8}$$

$$\alpha_l^n(0) = u_{0l}, , \quad l = 1, \ldots, n \tag{7.9}$$

with

$$\begin{aligned} a_{lj} &= \left(\nu \nabla \boldsymbol{v}_j^n, \nabla \boldsymbol{v}_l^n\right), \quad n_{ljk} = \left(\left(\boldsymbol{v}_j^n \cdot \nabla\right) \boldsymbol{v}_k^n, \boldsymbol{v}_l^n\right) = n\left(\boldsymbol{v}_j^n, \boldsymbol{v}_k^n, \boldsymbol{v}_l^n\right), \\ f_l &= \langle \boldsymbol{f}, \boldsymbol{v}_l^n \rangle_{V',V}, \quad u_{0l} = \left(\boldsymbol{u}_0, \boldsymbol{v}_l^n\right). \end{aligned}$$

The orthonormality of the basis functions is not essential but only simplifies the presentation. For non-orthonormal basis functions, the Gramian matrix (mass matrix) of V_{div}^n would appear at the first term of (7.8). This matrix is non-singular, thus multiplying with its inverse gives the same first term as in (7.8) and it leads to some changes in the other terms, e.g., see Temam (1984, p. 283). □

Lemma 7.12 (Unique Solvability of the Problem in the Finite-Dimensional Space) *Let $\Omega \in \mathbb{R}^d$, $d \in \{2, 3\}$, be a bounded Lipschitz domain. Let the regularity assumptions on $\boldsymbol{f}$ and $\boldsymbol{u}_0$ from Definitions 7.6 and 7.8 be satisfied. Then system (7.7)–(7.9) has a unique solution that is absolutely continuous in $[0, T]$. There hold the a priori estimates*

$$\sup_{t\in[0,T]} \|\boldsymbol{u}^n(t)\|_{L^2(\Omega)}^2 \le \|\boldsymbol{u}_0\|_{L^2(\Omega)}^2 + \frac{1}{\nu}\|\boldsymbol{f}\|_{L^2(0,T;V')}^2 \tag{7.10}$$

and

$$\|\boldsymbol{u}^n(T)\|_{L^2(\Omega)}^2 + \nu\,\|\nabla \boldsymbol{u}^n\|_{L^2(0,T;L^2(\Omega))}^2 \le \|\boldsymbol{u}_0\|_{L^2(\Omega)}^2 + \frac{1}{\nu}\|\boldsymbol{f}\|_{L^2(0,T;V')}^2, \tag{7.11}$$

which are both uniformly with respect to n. Hence,

$$\boldsymbol{u}^n \in L^2\left(0,T;V_{\text{div}}\right) \cap L^\infty\left(0,T;H_{\text{div}}(\Omega)\right).$$

Proof The proof of the lemma is based on the application of the theorem of Carathéodory, see Theorem A.50. Thus, one has to show a Lipschitz condition for the right-hand side of

$$\frac{d\alpha_l^n}{dt}(t) = F\left(\alpha_l^n\right), \quad t \in (0,T], \tag{7.12}$$

with $F \in L^2(0,T)$. If F would be continuous in $[0,T]$, one could apply the famous theorem of Peano, but for the given regularity, the theorem of Carathéodory has to be used.

The functions α_l^n appear linearly and quadratically on the right-hand side of (7.12). Hence, the Lipschitz condition is satisfied, since linear and quadratic functions are Lipschitz continuous. Consequently, the local existence and uniqueness of an absolutely continuous solution $\boldsymbol{u}^n(t,\boldsymbol{x})$ in some maximal interval $[0,t_n]$ with $0 < t_n \le T$ can be concluded from the theorem of Carathéodory. If $t_n < T$, then $\boldsymbol{u}^n(t)$ blows up as $t \to t_n$.

Next, the a priori estimates (7.10) and (7.11) will be proved which show that this situation cannot happen and therefore $t_n = T$.

Taking the solution $\boldsymbol{u}^n(t,\boldsymbol{x})$, which is now known to exist, as test function in (7.6) for some arbitrary $t \in (0,T)$, using the product rule for the first term

$$\begin{aligned}\frac{d}{dt}\left\|\boldsymbol{u}^n(t)\right\|_{L^2(\Omega)}^2 &= \frac{d}{dt}\left(\boldsymbol{u}^n(t),\boldsymbol{u}^n(t)\right) = \left(\partial_t\boldsymbol{u}^n(t),\boldsymbol{u}^n(t)\right) + \left(\boldsymbol{u}^n(t),\partial_t\boldsymbol{u}_n(t)\right)\\ &= 2\left(\partial_t\boldsymbol{u}^n(t),\boldsymbol{u}^n(t)\right),\end{aligned} \tag{7.13}$$

the skew-symmetry (6.28) of the convective term, the estimate of the dual pairing, and Young's inequality (A.5) yields

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\left\|\boldsymbol{u}^n(t)\right\|_{L^2(\Omega)}^2 + \nu\left\|\nabla\boldsymbol{u}^n(t)\right\|_{L^2(\Omega)}^2 &= \left\langle \boldsymbol{f}(t),\boldsymbol{u}^n(t)\right\rangle_{V',V} \le \left\|\boldsymbol{f}(t)\right\|_{V'}\left\|\nabla\boldsymbol{u}^n(t)\right\|_{L^2(\Omega)}\\ &\le \frac{1}{2\nu}\left\|\boldsymbol{f}(t)\right\|_{V'}^2 + \frac{\nu}{2}\left\|\nabla\boldsymbol{u}^n(t)\right\|_{L^2(\Omega)}^2,\end{aligned} \tag{7.14}$$

which gives

$$\frac{d}{dt}\left\|\boldsymbol{u}^n(t)\right\|_{L^2(\Omega)}^2 + \nu\left\|\nabla\boldsymbol{u}^n(t)\right\|_{L^2(\Omega)}^2 \le \frac{1}{\nu}\left\|\boldsymbol{f}(t)\right\|_{V'}^2. \tag{7.15}$$

Integrating this inequality in some time interval $[0, t]$ with arbitrary $t \leq T$ and applying the estimate for the $L^2(\Omega)$ projection (C.26) leads to

$$\|\boldsymbol{u}^n(t)\|_{L^2(\Omega)}^2 \leq \|\boldsymbol{u}_0^n\|_{L^2(\Omega)}^2 + \frac{1}{\nu}\int_0^t \|\boldsymbol{f}(\tau)\|_{V'}^2\, d\tau \leq \|\boldsymbol{u}_0\|_{L^2(\Omega)}^2 + \frac{1}{\nu}\|\boldsymbol{f}\|_{L^2(0,t;V')}^2 .$$

Taking the supremum in $[0, T]$ gives the a priori estimate (7.10) and since clearly $\boldsymbol{u}^n \in H_{\mathrm{div}}(\Omega)$, it follows that $\boldsymbol{u}^n \in L^\infty(0, T; H_{\mathrm{div}}(\Omega))$.

Integrating now (7.15) in $[0, T]$ leads in the same way as used for deriving the first estimate to

$$\|\boldsymbol{u}^n(T)\|_{L^2(\Omega)}^2 + \nu\,\|\nabla\boldsymbol{u}^n\|_{L^2(0,T;L^2(\Omega))}^2 \leq \|\boldsymbol{u}_0\|_{L^2(\Omega)}^2 + \frac{1}{\nu}\|\boldsymbol{f}\|_{L^2(0,T;V')}^2 ,$$

which is the a priori estimate (7.11). Since $\boldsymbol{u}^n \in V_{\mathrm{div}}$, one obtains in particular that $\boldsymbol{u}^n \in L^2(0, T; V_{\mathrm{div}})$. ■

Corollary 7.13 (Weak Convergence) *There is a subsequence* $\{\boldsymbol{u}^{n_l}\}_{l=1}^\infty$ *of* $\{\boldsymbol{u}^n\}_{n=1}^\infty$ *and a function* $\boldsymbol{u} \in L^2(0, T; V_{\mathrm{div}}) \cap L^\infty(0, T; H_{\mathrm{div}}(\Omega))$ *such that*

$$\begin{aligned} \boldsymbol{u}^{n_l} &\overset{*}{\rightharpoonup} \boldsymbol{u} \quad \text{in } L^\infty(0, T; H_{\mathrm{div}}(\Omega)), \\ \boldsymbol{u}^{n_l} &\rightharpoonup \boldsymbol{u} \quad \text{in } L^2(0, T; V_{\mathrm{div}}) \end{aligned}$$

as $l \to \infty$.

Proof It is known from Lemma 7.12 that the sequence $\{\boldsymbol{u}^n\}_{n=1}^\infty$ is bounded uniformly in $L^2(0, T; V_{\mathrm{div}})$. This space is a Hilbert space, thus in particular a reflexive Banach space, such that the existence of a weakly convergence subsequence to some element $\boldsymbol{u}_1 \in L^2(0, T; V_{\mathrm{div}})$ follows, see Remark A.58. The limit is unique.

The space $L^1(0, T; H_{\mathrm{div}}(\Omega))$ is a separable Banach space and its dual space is $L^\infty(0, T; H_{\mathrm{div}}(\Omega))$. From Lemma 7.12, it is known that $\{\boldsymbol{u}^n\}_{n=1}^\infty$ is bounded uniformly in the dual space, such that the existence of a weakly* convergent subsequence to some function $\boldsymbol{u}_2 \in L^\infty(0, T; H_{\mathrm{div}}(\Omega))$, follows, see Remark A.58. Also this limit is unique.

Since

$$\{\boldsymbol{u}^n\}_{n=1}^\infty \in L^2(0, T; V_{\mathrm{div}}) \cap L^\infty(0, T; H_{\mathrm{div}}(\Omega))$$

and both spaces are complete, it follows that

$$\boldsymbol{u}_1, \boldsymbol{u}_2 \in L^2(0, T; V_{\mathrm{div}}) \cap L^\infty(0, T; H_{\mathrm{div}}(\Omega)) .$$

The uniqueness of the limits gives finally that $\boldsymbol{u}_1 = \boldsymbol{u}_2 = \boldsymbol{u}$. ■

Remark 7.14 (Consequences of the Weak Convergence, Convergence of the Linear Terms) For simplicity of notation, the subsequence will be denoted again by $\{\boldsymbol{u}^n\}_{n=1}^{\infty}$.

Considering the solution of the Galerkin problem (7.6), taking an arbitrary $\boldsymbol{v} \in V_{\text{div}}$, and letting $\phi \in C_0^\infty((0,T))$, then integration by parts gives

$$\int_0^T (\partial_t \boldsymbol{u}^n(t), \boldsymbol{v})\, \phi(t)\, dt = -\int_0^T (\boldsymbol{u}^n(t), \boldsymbol{v})\, \frac{d}{dt}\phi(t)\, dt.$$

Since $\boldsymbol{u}^n \overset{*}{\rightharpoonup} \boldsymbol{u}$ in $L^\infty(0,T;H_{\text{div}}(\Omega))$ and $\frac{d}{dt}\phi \in L^1(0,T)$, one gets

$$\begin{aligned}\lim_{n\to\infty} -\int_0^T (\boldsymbol{u}^n(t), \boldsymbol{v})\, \frac{d}{dt}\phi(t)\, dt &= \int_0^T (\boldsymbol{u}(t), \boldsymbol{v})\, \frac{d}{dt}\phi(t)\, dt \\ &= \int_0^T (\partial_t \boldsymbol{u}(t), \boldsymbol{v})\, \phi(t)\, dt\end{aligned}$$

for all $\boldsymbol{v} \in V_{\text{div}}$. Hence, $\boldsymbol{u}$ satisfies the weak form of the first term of (7.3).

For the viscous term, it follows from $\boldsymbol{u}^n \rightharpoonup \boldsymbol{u}$ in $L^2(0,T;V_{\text{div}})$ and $\phi \in L^2(0,T)$ that for all $\boldsymbol{v} \in V_{\text{div}}$ and $\phi \in C_0^\infty((0,T))$

$$\lim_{n\to\infty} \int_0^T (\nu \nabla \boldsymbol{u}^n, \nabla \boldsymbol{v})\, \phi(t)\, dt = \int_0^T (\nu \nabla \boldsymbol{u}, \nabla \boldsymbol{v})\, \phi(t)\, dt, \tag{7.16}$$

which is the weak form of the viscous term in (7.3).

In order to show that $\boldsymbol{u}$ is a weak solution of the Navier–Stokes equations, one has to show the convergence of the nonlinear convective term

$$n\left(\boldsymbol{u}^h, \boldsymbol{u}^n, \boldsymbol{v}\right) \to n\left(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}\right).$$

There are different ways to prove this property, e.g., using a Fourier transform as in Temam (1984, p. 285) or a special Friedrichs inequality as in Galdi (2000). Below, the approach used in Girault and Raviart (1979, p. 161) will be sketched. □

Lemma 7.15 (Estimate of the Difference of Nonlinear Convective Terms) *Let* $\boldsymbol{u}, \overline{\boldsymbol{u}} \in L^2(0,T;V_{\text{div}}) \cap L^\infty(0,T;H_{\text{div}}(\Omega))$, $\boldsymbol{v} \in V_{\text{div}}$, *and* $r \in [1, 4/3)$, *then*

$$\begin{aligned}&\int_0^T |n(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}) - n(\overline{\boldsymbol{u}},\overline{\boldsymbol{u}},\boldsymbol{v})|^r\, dt \\ &\le C\, \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}^r \left(\|\boldsymbol{u}\|_{L^\infty(0,T;L^2(\Omega))}^{r/4} + \|\overline{\boldsymbol{u}}\|_{L^\infty(0,T;L^2(\Omega))}^{r/4}\right) \\ &\quad \times \|\boldsymbol{u}-\overline{\boldsymbol{u}}\|_{L^{r/(4-3r)}(0,T;L^2(\Omega))}^{r/4} \left(\|\nabla \boldsymbol{u}\|_{L^2(0,T;L^2(\Omega))}^{3r/2} + \|\nabla \overline{\boldsymbol{u}}\|_{L^2(0,T;L^2(\Omega))}^{3r/2}\right).\end{aligned} \tag{7.17}$$

Proof By adding and subtracting the same term, applying the triangle inequality, and using the skew-symmetry of the nonlinear convective term (6.24), one obtains

$$\begin{aligned}|n(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v})-n(\overline{\boldsymbol{u}},\overline{\boldsymbol{u}},\boldsymbol{v})| &= |n(\boldsymbol{u}-\overline{\boldsymbol{u}},\boldsymbol{u},\boldsymbol{v})-n(\overline{\boldsymbol{u}},\overline{\boldsymbol{u}}-\boldsymbol{u},\boldsymbol{v})| \\ &\le |n(\boldsymbol{u}-\overline{\boldsymbol{u}},\boldsymbol{u},\boldsymbol{v})|+|n(\overline{\boldsymbol{u}},\overline{\boldsymbol{u}}-\boldsymbol{u},\boldsymbol{v})| \\ &= |n(\boldsymbol{u}-\overline{\boldsymbol{u}},\boldsymbol{v},\boldsymbol{u})|+|n(\overline{\boldsymbol{u}},\boldsymbol{v},\overline{\boldsymbol{u}}-\boldsymbol{u})|\,.\end{aligned}$$

The goal of the proof is to derive an estimated where the norm of the difference $\boldsymbol{u}-\overline{\boldsymbol{u}}$ is as weak as possible.

The next steps of the estimate consist in applying the generalized Hölder inequality (6.33) with $p=r=1/4$ and $q=2$, using the Sobolev imbedding (A.15) with $m=3/4$, $p=2$, $q=4$, and using the interpolation estimate for Sobolev spaces (A.13)

$$\begin{aligned}&|n(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v})-n(\overline{\boldsymbol{u}},\overline{\boldsymbol{u}},\boldsymbol{v})| \\ &\le \|\boldsymbol{u}-\overline{\boldsymbol{u}}\|_{L^4(\Omega)}\,\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}\,\|\boldsymbol{u}\|_{L^4(\Omega)}+\|\overline{\boldsymbol{u}}\|_{L^4(\Omega)}\,\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}\,\|\overline{\boldsymbol{u}}-\boldsymbol{u}\|_{L^4(\Omega)} \\ &= \|\nabla\boldsymbol{v}\|_{L^2(\Omega)}\,\|\boldsymbol{u}-\overline{\boldsymbol{u}}\|_{L^4(\Omega)}\left(\|\boldsymbol{u}\|_{L^4(\Omega)}+\|\overline{\boldsymbol{u}}\|_{L^4(\Omega)}\right) \\ &\le C\,\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}\,\|\boldsymbol{u}-\overline{\boldsymbol{u}}\|_{H^{3/4}(\Omega)}\left(\|\boldsymbol{u}\|_{H^{3/4}(\Omega)}+\|\overline{\boldsymbol{u}}\|_{H^{3/4}(\Omega)}\right) \\ &\le C\,\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}\,\|\boldsymbol{u}-\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{1/4}\,\|\nabla\left(\boldsymbol{u}-\overline{\boldsymbol{u}}\right)\|_{L^2(\Omega)}^{3/4} \\ &\quad\times\left(\|\boldsymbol{u}\|_{L^2(\Omega)}^{1/4}\,\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^{3/4}+\|\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{1/4}\,\|\nabla\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{3/4}\right).\end{aligned}$$

Note that in the application of the generalized Hölder inequality, the maximal regularity for $\boldsymbol{v}$ was used and the other two terms are treated in the same way. Note also that the application of Poincaré's inequality would lead to a strong norm for the difference $\boldsymbol{u}-\overline{\boldsymbol{u}}$.

Let $r\in[1,4/3)$, then it follows that

$$\begin{aligned}&\int_0^T|n(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v})-n(\overline{\boldsymbol{u}},\overline{\boldsymbol{u}},\boldsymbol{v})|^r\,dt \\ &\le C\,\|\nabla\boldsymbol{v}\|_{L^2(\Omega)}^r\int_0^T\Big[\,\|\boldsymbol{u}-\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{r/4}\,\|\nabla\left(\boldsymbol{u}-\overline{\boldsymbol{u}}\right)\|_{L^2(\Omega)}^{3r/4} \\ &\quad\times\left(\|\boldsymbol{u}\|_{L^2(\Omega)}^{1/4}\,\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^{3/4}+\|\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{1/4}\,\|\nabla\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{3/4}\right)^r\Big]\,dt.\end{aligned}\tag{7.18}$$

Taking the supremum in $(0,T)$ of two terms, using

$$(ab+cd)\le(a+c)(b+d),\quad a,b,c,d\ge0,$$

and applying $(a+b)^n \le C(a^n + b^n)$ for $a, b \ge 0$, $n \ge 1$, gives for the integral

$$\begin{aligned}
&\int_0^T \Big[\|\boldsymbol{u}-\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{r/4} \|\nabla(\boldsymbol{u}-\overline{\boldsymbol{u}})\|_{L^2(\Omega)}^{3r/4} \\
&\quad \times \Big(\|\boldsymbol{u}\|_{L^2(\Omega)}^{1/4} \|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^{3/4} + \|\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{1/4} \|\nabla \overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{3/4} \Big)^r \Big] \, dt \\
&\le \int_0^T \Big[\|\boldsymbol{u}-\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{r/4} \|\nabla(\boldsymbol{u}-\overline{\boldsymbol{u}})\|_{L^2(\Omega)}^{3r/4} \\
&\quad \times \Big(\|\boldsymbol{u}\|_{L^\infty(0,T;L^2(\Omega))}^{1/4} \|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^{3/4} + \|\overline{\boldsymbol{u}}\|_{L^\infty(0,T;L^2(\Omega))}^{1/4} \|\nabla \overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{3/4} \Big)^r \Big] \, dt \\
&\le \int_0^T \Big[\|\boldsymbol{u}-\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{r/4} \|\nabla(\boldsymbol{u}-\overline{\boldsymbol{u}})\|_{L^2(\Omega)}^{3r/4} \Big(\|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^{3/4} + \|\nabla \overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{3/4} \Big)^r \\
&\quad \times \Big(\|\boldsymbol{u}\|_{L^\infty(0,T;L^2(\Omega))}^{1/4} + \|\overline{\boldsymbol{u}}\|_{L^\infty(0,T;L^2(\Omega))}^{1/4} \Big)^r \Big] \, dt \\
&\le C \Big(\|\boldsymbol{u}\|_{L^\infty(0,T;L^2(\Omega))}^{r/4} + \|\overline{\boldsymbol{u}}\|_{L^\infty(0,T;L^2(\Omega))}^{r/4} \Big) \\
&\quad \times \int_0^T \Big[\|\boldsymbol{u}-\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{r/4} \|\nabla(\boldsymbol{u}-\overline{\boldsymbol{u}})\|_{L^2(\Omega)}^{3r/4} \Big(\|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^{3r/4} + \|\nabla \overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{3r/4} \Big) \Big] \, dt.
\end{aligned} \tag{7.19}$$

Concentrating again on the integral, using the triangle inequality (in the case of Lebesgue spaces also called Minkowski's inequality),

$$(a+b)^q \le a^q + b^q \quad a, b \ge 0, q \in (0,1),$$

which follows from the left inequality of (A.4), and Young's inequality (A.5) yields

$$\begin{aligned}
&\int_0^T \Big[\|\boldsymbol{u}-\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{r/4} \|\nabla(\boldsymbol{u}-\overline{\boldsymbol{u}})\|_{L^2(\Omega)}^{3r/4} \Big(\|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^{3r/4} + \|\nabla \overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{3r/4} \Big) \Big] \, dt \\
&\le \int_0^T \Big[\|\boldsymbol{u}-\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{r/4} \Big(\|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^{3r/4} + \|\nabla \overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{3r/4} \Big) \Big(\|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^{3r/4} + \|\nabla \overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{3r/4} \Big) \Big] \, dt \\
&\le C \int_0^T \|\boldsymbol{u}-\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{r/4} \Big(\|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^{3r/2} + \|\nabla \overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{3r/2} \Big) \, dt.
\end{aligned} \tag{7.20}$$

Note that the integral is well defined for $r \le 4/3$ since the first term is in $L^\infty(0,T)$ by assumption and the norms with the gradient are in $L^2(0,T)$.

Next, Hölder's inequality (A.9) will be applied. Since $\boldsymbol{u}, \overline{\boldsymbol{u}} \in L^2(0,T;V_{\text{div}})$, the terms in the parentheses should get the power 2. Therefore, one chooses in (A.9) the values $p = 4/(3r) > 1$, $q = 4/(4-3r)$ and one obtains

$$\begin{aligned}
&\int_0^T \left[\|\boldsymbol{u}-\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{r/4} \|\nabla(\boldsymbol{u}-\overline{\boldsymbol{u}})\|_{L^2(\Omega)}^{3r/4} \left(\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^{3r/4} + \|\nabla\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{3r/4} \right) \right] dt \\
&\le \left(\int_0^T \|\boldsymbol{u}-\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^{r/(4-3r)} \, dt \right)^{(4-3r)/4} \\
&\quad \times \left[\left(\int_0^T \|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^2 \, dt \right)^{3r/4} + \left(\int_0^T \|\nabla\overline{\boldsymbol{u}}\|_{L^2(\Omega)}^2 \, dt \right)^{3r/4} \right].
\end{aligned}$$

Inserting this estimate in (7.18), (7.19), and (7.20) gives estimate (7.17). ■

Lemma 7.16 (Convergence of the Nonlinear Convective Term) *It holds for all* $\boldsymbol{v} \in V_{\text{div}}$ *and* $\phi \in C_0^\infty((0,T))$

$$\lim_{n\to\infty} \int_0^T n\left(\boldsymbol{u}^n(t), \boldsymbol{u}^n(t), \boldsymbol{v}\right) \phi(t) \, dt = \int_0^T n\left(\boldsymbol{u}(t), \boldsymbol{u}(t), \boldsymbol{v}\right) \phi(t) \, dt \tag{7.21}$$

in $L^2(0,T;V_{\text{div}})$.

Proof Only a sketch of the proof will be presented.

In the first step, one proves that

$$\boldsymbol{u}^n \in H^\sigma(0,T;H_{\text{div}}(\Omega)) \quad \text{with } \sigma \in \left(0, \frac{1}{4}\right),$$

where the space is equipped with the norm

$$\|\boldsymbol{u}^n\|_{H^\sigma(0,T;H_{\text{div}}(\Omega))} = \int_0^T \int_0^T |t-s|^{-(1+2\sigma)} \|\boldsymbol{u}^n(t) - \boldsymbol{u}^n(s)\|_{L^2(\Omega)} \, dsdt.$$

The proof of this property is quite lengthy and it is referred to the literature, e.g., Girault and Raviart (1979, Chap. V, Lemma 1.7).

Applying, e.g., the generalized theorem of Lions–Aubin, it is shown next that

$$L^2(0,T;V_{\text{div}}) \cap H^\sigma(0,T;H_{\text{div}}(\Omega))$$

is compactly imbedded into $L^2(0,T;H_{\text{div}}(\Omega))$, e.g., see Girault and Raviart (1979, Chap. V, Lemma 1.3) for a direct proof of this property. Because $\{\boldsymbol{u}^n\}_{n=1}^\infty$ is bounded in $L^2(0,T;V_{\text{div}}) \cap H^\sigma(0,T;H_{\text{div}}(\Omega))$, it follows from the compactness of the imbedding that there is a subsequence that is a Cauchy sequence in $L^2(0,T;H_{\text{div}}(\Omega))$. Since $L^2(0,T;H_{\text{div}}(\Omega))$ is complete, this subsequence converges strongly to some function $\boldsymbol{u} \in L^2(0,T;H_{\text{div}}(\Omega))$. Because of the uniqueness of the limit and because this limit is contained in $L^2(0,T;V_{\text{div}})$, it has to be the same limit as the limit from Corollary 7.13.

Taking $\overline{\boldsymbol{u}} = \boldsymbol{u}^n$ and $r = 8/7$ in (7.17) gives the term

$$\|\boldsymbol{u} - \boldsymbol{u}^n\|_{L^{r/(4-3r)}(0,T;L^2(\Omega))}^{r/4} = \|\boldsymbol{u} - \boldsymbol{u}^n\|_{L^2(0,T;L^2(\Omega))}^{2/7},$$

which converges to zero as $n \to \infty$. Thus, one gets from (7.17) that

$$\lim_{n\to\infty} \int_0^T |n(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}) - n(\boldsymbol{u}^n, \boldsymbol{u}^n, \boldsymbol{v})|^{8/7} \, dt = 0.$$

Since this expression converges for the power $8/7$, it converges also for the power one, such that one obtains

$$\lim_{n\to\infty} \int_0^T |n(\boldsymbol{u}^n, \boldsymbol{u}^n, \boldsymbol{v}) - n(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}))| \, dt = 0.$$

Finally, it follows that

$$\begin{aligned} &\lim_{n\to\infty} \left| \int_0^T (n(\boldsymbol{u}^n, \boldsymbol{u}^n, \boldsymbol{v}) - n(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}))) \, \phi(t) \, dt \right| \\ &\le \|\phi\|_{L^\infty((0,T))} \lim_{n\to\infty} \int_0^T |n(\boldsymbol{u}^n, \boldsymbol{u}^n, \boldsymbol{v}) - n(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}))| \, dt = 0, \end{aligned}$$

which is the statement of the lemma. ∎

Lemma 7.17 (Satisfaction of the Initial Condition) *The limit* $\boldsymbol{u}$ *from Corollary 7.13 and Lemma 7.16 satisfies the initial condition.*

Proof Let $\boldsymbol{v}^n \in V_{\text{div}}^n$ be arbitrary, then the fundamental theorem of calculus and the product rule yields

$$\begin{aligned} -(\boldsymbol{u}(0), \boldsymbol{v}^n) &= (\boldsymbol{u}(t), \boldsymbol{v}^n) \frac{T-t}{T}\bigg|_{t=0}^{T} = \int_0^T \frac{d}{dt} \left((\boldsymbol{u}(t), \boldsymbol{v}^n) \frac{T-t}{T} \right) dt \\ &= \int_0^T \left(\frac{d}{dt} (\boldsymbol{u}, \boldsymbol{v}^n) \right) \frac{T-t}{T} \, dt - \frac{1}{T} \int_0^T (\boldsymbol{u}, \boldsymbol{v}^n) \, dt. \end{aligned}$$

Inserting for the time derivative equation (7.3) and replacing the term with the exterior force by (7.6) gives

$$\begin{aligned} &-(\boldsymbol{u}(0), \boldsymbol{v}^n) \\ &= \int_0^T (\langle \boldsymbol{f}, \boldsymbol{v}^n \rangle_{V',V} - (\nu \nabla \boldsymbol{u}, \nabla \boldsymbol{v}^n) - n(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}^n)) \frac{T-t}{T} \, dt - \frac{1}{T} \int_0^T (\boldsymbol{u}, \boldsymbol{v}^n) \, dt \end{aligned}$$

$$
\begin{aligned}
&= \int_0^T \left((\nu \nabla \boldsymbol{u}^n, \nabla \boldsymbol{v}^n) - (\nu \nabla \boldsymbol{u}, \nabla \boldsymbol{v}^n) \right) \frac{T-t}{T} \, dt \\
&\quad + \int_0^T \left(n\left(\boldsymbol{u}^n, \boldsymbol{u}^n, \boldsymbol{v}^n\right) - n\left(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}^n\right) \right) \frac{T-t}{T} \, dt \\
&\quad + \int_0^T \left(\partial_t \boldsymbol{u}^n, \boldsymbol{v}^n\right) \frac{T-t}{T} \, dt - \frac{1}{T} \int_0^T \left(\boldsymbol{u}, \boldsymbol{v}^n\right) \, dt.
\end{aligned}
$$

Now, one considers $n \to \infty$. From (7.16) and (7.21), it follows that the first two terms vanish. One obtains, using integration by parts in time and $\boldsymbol{u}^n \overset{*}{\rightharpoonup} \boldsymbol{u}$ in $L^\infty\left(0, T; H_{\text{div}}(\Omega)\right)$

$$
\begin{aligned}
&- (\boldsymbol{u}(0), \boldsymbol{v}) \\
&= \lim_{n\to\infty} - \left(\boldsymbol{u}(0), \boldsymbol{v}^n\right) \\
&= \lim_{n\to\infty} \int_0^T \left(\partial_t \boldsymbol{u}^n, \boldsymbol{v}^n\right) \frac{T-t}{T} \, dt - \frac{1}{T} \int_0^T \left(\boldsymbol{u}, \boldsymbol{v}^n\right) \, dt \\
&= \lim_{n\to\infty} \left(\boldsymbol{u}^n, \boldsymbol{v}^n\right) \frac{T-t}{T} \bigg|_{t=0}^T + \frac{1}{T} \lim_{n\to\infty} \left(\int_0^T \left(\boldsymbol{u}^n, \boldsymbol{v}^n\right) \, dt - \int_0^T \left(\boldsymbol{u}, \boldsymbol{v}^n\right) \, dt \right) \\
&= - \lim_{n\to\infty} \left(\boldsymbol{u}^n(0), \boldsymbol{v}^n\right) \\
&= - \lim_{n\to\infty} \left(\boldsymbol{u}_0^n, \boldsymbol{v}^n\right).
\end{aligned}
$$

Hence, the limit function $\boldsymbol{u}$ satisfies the initial condition. ■

Theorem 7.18 (Existence of a Weak Solution in the Sense of Leray–Hopf) *Let* $\boldsymbol{f} \in L^2\left(0, T; V'\right)$ *and* $\boldsymbol{u}_0 \in H_{\text{div}}(\Omega)$*, then there exists a weak solution*

$$
\boldsymbol{u} \in L^2\left(0, T; V_{\text{div}}\right) \cap L^\infty\left(0, T; H_{\text{div}}(\Omega)\right)
$$

of (7.3) *or* (7.4) *in the sense of Leray–Hopf.*

Proof From Remark 7.14 and Lemma 7.16, it follows that there is a subsequence of $\{\boldsymbol{u}^n\}_{n=1}^\infty$ whose limit $\boldsymbol{u}$ satisfies the weak equation. In Lemma 7.17 it was proved that also the initial condition is satisfied from $\boldsymbol{u}$. ■

Lemma 7.19 (Regularity of the Time Derivative) *Let* $\boldsymbol{u} \in L^2\left(0, T; V_{\text{div}}\right) \cap L^\infty\left(0, T; H_{\text{div}}(\Omega)\right)$ *be the Leray–Hopf weak solution of* (7.3)*, then*

$$
\partial_t \boldsymbol{u} \in \begin{cases} L^2\left(0, T; V'\right) & \text{if } d = 2, \\ L^{4/3}\left(0, T; V'\right) & \text{if } d = 3. \end{cases} \tag{7.22}
$$

Proof By definition, it is

$$\|\partial_t \boldsymbol{u}\|^p_{L^p(0,T;V')} = \int_0^T \left(\sup_{\substack{\boldsymbol{v}\in V \\ \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}=1}} \langle \partial_t \boldsymbol{u}, \boldsymbol{v}\rangle_{V',V} \right)^p dt.$$

Replacing the term in the parentheses with the weak form of the Navier–Stokes equations (7.4), applying the estimate for the dual pairing and the Cauchy–Schwarz inequality (A.10), and using (A.4) yields

$$\begin{aligned}
&\|\partial_t \boldsymbol{u}\|^p_{L^p(0,T;V')} \\
&= \int_0^T \left(\sup_{\substack{\boldsymbol{v}\in V \\ \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}=1}} \left(\langle \boldsymbol{f}, \boldsymbol{v}\rangle_{V',V} - \nu\left(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}\right) - n\left(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}\right)\right) \right)^p dt \\
&\le \int_0^T \left(\sup_{\substack{\boldsymbol{v}\in V \\ \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}=1}} \left(\|\boldsymbol{f}\|_{V'} \|\nabla \boldsymbol{v}\|_{L^2(\Omega)} + \nu \|\nabla \boldsymbol{u}\|_{L^2(\Omega)} \|\nabla \boldsymbol{v}\|_{L^2(\Omega)} + |n\left(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}\right)|\right) \right)^p dt \\
&\le C \int_0^T \|\boldsymbol{f}\|^p_{V'} + \|\nabla \boldsymbol{u}\|^p_{L^2(\Omega)} + \left(\sup_{\substack{\boldsymbol{v}\in V \\ \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}=1}} |n\left(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}\right)| \right)^p dt. \qquad (7.23)
\end{aligned}$$

The first term is bounded for $p \le 2$ since one obtains with Hölder's inequality (A.9)

$$\int_0^T \|\boldsymbol{f}\|^p_{V'}\ dt \le \left(\int_0^T \|\boldsymbol{f}\|^2_{V'}\ dt\right)^{p/2} \left(\int_0^T 1^{2/(2-p)}\ dt\right)^{(2-p)/2} = C\, \|\boldsymbol{f}\|^{p/2}_{L^2(0,T;V')} < \infty,$$

by the regularity assumption on $\boldsymbol{f}$. In the same way, using the regularity assumption (7.5) for a weak solution, one finds that the second term of (7.23) is finite.

For the third term of (7.23), one gets with (6.24) and (6.33) with $p = r = 4$

$$\begin{aligned}
\int_0^T \left(\sup_{\substack{\boldsymbol{v}\in V \\ \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}=1}} |n\left(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}\right)| \right)^p dt &= \int_0^T \left(\sup_{\substack{\boldsymbol{v}\in V \\ \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}=1}} |-n\left(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{u}\right)| \right)^p dt \\
&\le \int_0^T \left(\sup_{\substack{\boldsymbol{v}\in V \\ \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}=1}} \|\boldsymbol{u}\|^2_{L^4(\Omega)} \|\nabla \boldsymbol{v}\|_{L^2(\Omega)} \right)^p dt \\
&= \int_0^T \|\boldsymbol{u}\|^{2p}_{L^4(\Omega)}\ dt. \qquad (7.24)
\end{aligned}$$

The question, for which values of p the right-hand side of (7.24) is finite will be studied separately for two and three dimensions.

Two-dimensional Case One has by the Sobolev imbedding $H^{1/2}(\Omega) \to L^4(\Omega)$, see (A.15), the interpolation theorem between Sobolev spaces (A.13), and Poincaré's inequality (A.12)

$$\|\boldsymbol{u}\|_{L^4(\Omega)}^4 \le C\|\boldsymbol{u}\|_{H^{1/2}(\Omega)}^4 \le C\|\boldsymbol{u}\|_{L^2(\Omega)}^2\|\boldsymbol{u}\|_{H^1(\Omega)}^2 \le C\|\boldsymbol{u}\|_{L^2(\Omega)}^2\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^2.$$

It follows that

$$\begin{aligned}\int_0^t \|\boldsymbol{u}(\tau)\|_{L^4(\Omega)}^4\,d\tau &\le C\int_0^t \|\boldsymbol{u}(\tau)\|_{L^2(\Omega)}^2\|\nabla\boldsymbol{u}(\tau)\|_{L^2(\Omega)}^2\,d\tau\\ &\le C\|\boldsymbol{u}\|_{L^\infty((0,T);L^2(\Omega))}^2\int_0^t\|\nabla\boldsymbol{u}(\tau)\|_{L^2(\Omega)}^2\,d\tau\\ &= C\|\boldsymbol{u}\|_{L^\infty((0,T);L^2(\Omega))}^2\|\boldsymbol{u}\|_{L^2((0,T);V)}^2 < \infty, \end{aligned} \qquad (7.25)$$

by the regularity properties of a weak solution. Together with the results for the first two terms of (7.23), the lemma is proved for the two-dimensional case.

Three-dimensional Case In three dimensions, the Sobolev imbedding $H^{3/4}(\Omega) \to L^4(\Omega)$ holds, see (A.15). Hence, one obtains with the interpolation theorem between Sobolev spaces (A.13) and Poincaré's inequality

$$\begin{aligned}\int_0^T \|\boldsymbol{u}\|_{L^4(\Omega)}^{2p}\,dt \le C\int_0^T\|\boldsymbol{u}\|_{H^{3/4}(\Omega)}^{2p}\,dt &\le C\int_0^T\|\boldsymbol{u}\|_{L^2(\Omega)}^{p/2}\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^{3p/2}\,dt\\ &\le C\|\boldsymbol{u}\|_{L^\infty((0,T);L^2(\Omega))}^{p/2}\int_0^T\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^{3p/2}\,dt. \end{aligned} \qquad (7.26)$$

By the regularity assumption (7.5) on a weak solution, the first factor is finite and the second factor is finite for $p = 3/4$. This result finishes the proof of the lemma for the three-dimensional case. ■

Definition 7.20 (Energy Inequality) A weak solution $\boldsymbol{u}$ is said to satisfy the energy inequality if

$$\begin{aligned}&\|\boldsymbol{u}(t)\|_{L^2(\Omega)}^2 + 2\nu\int_0^t\|\nabla\boldsymbol{u}(\tau)\|_{L^2(\Omega)}^2\,d\tau\\ &\le \|\boldsymbol{u}(0)\|_{L^2(\Omega)}^2 + 2\int_0^t\langle\boldsymbol{f},\boldsymbol{u}\rangle_{V'V}(\tau)\,d\tau \end{aligned} \qquad (7.27)$$

for all $t \in [0,T]$. The first term on the left-hand side is (twice of) the kinetic energy and the second term is the energy dissipation due to the viscosity of the fluid. □

Lemma 7.21 (Energy Inequality, Stability of the Solution) *Let $d = 2$ and let $\boldsymbol{u}$ be any weak solution of* (7.3) *or* (7.4). *Then, this solution satisfies an energy equality. In addition, the following stability bound for the velocity holds*

$$\|\boldsymbol{u}(t)\|_{L^2(\Omega)}^2 + \nu \|\nabla\boldsymbol{u}\|_{L^2(0,t;L^2(\Omega))}^2 \le \|\boldsymbol{u}(0)\|_{L^2(\Omega)}^2 + \frac{1}{\nu}\|\boldsymbol{f}\|_{L^2(0,t;H^{-1}(\Omega))}^2 \qquad (7.28)$$

for all $t \in [0, T]$.

If $d = 3$ and $\boldsymbol{u}$ is the weak solution of (7.3) *or* (7.4) *in the sense of Leray–Hopf, then this solution satisfies the energy inequality* (7.27) *and a stability estimate of form* (7.28).

Proof The proof is different for two and three dimensions.

Two-dimensional Case In this case, one can apply the usual approach, i.e., one can take the weak solution $\boldsymbol{u}$ as a test function. One obtains from (7.3), (7.13), and Lemma 6.10

$$\frac{1}{2}\frac{d}{dt}\|\boldsymbol{u}\|_{L^2(\Omega)}^2 + \nu\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^2 = \langle \boldsymbol{f}, \boldsymbol{u}\rangle_{V',V}.$$

Integrating in $(0, t)$ yields

$$\frac{1}{2}\|\boldsymbol{u}(t)\|_{L^2(\Omega)}^2 + \nu\|\nabla\boldsymbol{u}\|_{L^2(0,t;L^2(\Omega))}^2 = \frac{1}{2}\|\boldsymbol{u}(0)\|_{L^2(\Omega)}^2 + \int_0^t \langle \boldsymbol{f}, \boldsymbol{u}\rangle_{V',V}(\tau)\, d\tau,$$

which is just the energy equality. Then, using the inequality for the dual pairing and Young's inequality (A.5) gives

$$\begin{aligned}&\frac{1}{2}\|\boldsymbol{u}(t)\|_{L^2(\Omega)}^2 + \nu\|\nabla\boldsymbol{u}\|_{L^2(0,t;L^2(\Omega))}^2 \\ &\le \frac{1}{2}\|\boldsymbol{u}(0)\|_{L^2(\Omega)}^2 + \frac{1}{2\nu}\|\boldsymbol{f}\|_{L^2(0,t;V')}^2 + \frac{\nu}{2}\|\nabla\boldsymbol{u}\|_{L^2(0,t;L^2(\Omega))}^2,\end{aligned}$$

which leads directly to (7.28).

Three-dimensional Case The difficulty in three dimensions with the usual approach is that the term

$$\int_0^t (\partial_t\boldsymbol{u}, \boldsymbol{u})(\tau)\, d\tau$$

might not be well defined, since $\partial_t\boldsymbol{u}$ is contained in $L^{4/3}(0, t; V')$, see (7.22), but not in $L^2(0, t; V')$. Thus, with respect to time, the term in the integral might not be in $L^1(0, t)$ such that the integral might not be well defined.

Starting point for proving the energy inequality is the first line of (7.14), i.e.,

$$\frac{1}{2}\frac{d}{dt}\|\boldsymbol{u}^n(t)\|_{L^2(\Omega)}^2 + \nu\,\|\nabla\boldsymbol{u}^n(t)\|_{L^2(\Omega)}^2 = \langle\boldsymbol{f}(t),\boldsymbol{u}^n(t)\rangle_{V',V}.$$

Integration in $(0,t)$ gives

$$\|\boldsymbol{u}^n(t)\|_{L^2(\Omega)}^2 + 2\nu\,\|\nabla\boldsymbol{u}^n\|_{L^2(0,t;L^2(\Omega))}^2 = \|\boldsymbol{u}^n(0)\|_{L^2(\Omega)}^2 + 2\int_0^t \langle\boldsymbol{f},\boldsymbol{u}^n\rangle_{V',V}(\tau)\,d\tau. \tag{7.29}$$

Since there is a subsequence of $\{\boldsymbol{u}^n\}_{n=1}^{\infty}$, which will be denoted for simplicity with the same symbol, that converges weakly to $\boldsymbol{u}$ in $L^2(0,T;V_{\mathrm{div}})$ and weakly* to $\boldsymbol{u}$ in $L^\infty(0,T;H_{\mathrm{div}})$, see Corollary 7.13, it follows that, e.g., see Evans (2010, p. 723),

$$\|\boldsymbol{u}\|_{L^\infty(0,t;L^2(\Omega))} \le \liminf_{n\to\infty}\|\boldsymbol{u}^n\|_{L^\infty(0,t;L^2(\Omega))},$$
$$\|\boldsymbol{u}\|_{L^2(0,t;V)} \le \liminf_{n\to\infty}\|\boldsymbol{u}^n\|_{L^2(0,t;V)}$$

for $t\in[0,T]$. Taking $n\to\infty$ in (7.29), using these estimates, noting that $\|\boldsymbol{u}(t)\|_{L^2(\Omega)} \le \|\boldsymbol{u}\|_{L^\infty(0,t;L^2(\Omega))}$, using that the limit satisfies the initial condition in the sense of $L^2(\Omega)$, see Lemma 7.17, and using the weak convergence in $L^2(0,t;V_{\mathrm{div}})$ leads to

$$\begin{aligned}
&\|\boldsymbol{u}(t)\|_{L^2(\Omega)}^2 + 2\nu\,\|\nabla\boldsymbol{u}\|_{L^2(0,t;L^2(\Omega))}^2\\
&\le \lim_{n\to\infty}\|\boldsymbol{u}^n(0)\|_{L^2(\Omega)}^2 + 2\lim_{n\to\infty}\int_0^t\langle\boldsymbol{f},\boldsymbol{u}^n\rangle_{V',V}(\tau)\,d\tau\\
&= \|\boldsymbol{u}(0)\|_{L^2(\Omega)}^2 + 2\int_0^t\langle\boldsymbol{f},\boldsymbol{u}\rangle_{V',V}(\tau)\,d\tau,
\end{aligned}$$

which is just the energy inequality (7.27). Estimate (7.28) is now proved analogously to the two-dimensional case by using the estimate for the dual pairing and Young's inequality. ■

Lemma 7.22 (Uniqueness of a Weak Solution with Stronger Regularity Assumptions on the Solution) *There is not more than one weak solution that satisfies* $\boldsymbol{u}\in L^8\left(0,T;L^4(\Omega)\right)$.

Proof First of all, one finds, with the assumed regularity, that it is also in three dimensions $\partial_t\boldsymbol{u}\in L^2(0,T;V')$. The proof of this statement follows the lines of the proof of Lemma 7.19. Starting with (7.26), using the Sobolev imbedding $H^{3/4}(\Omega)\to L^4(\Omega)$, see (A.15), the interpolation in Sobolev spaces (A.13), and

Hölder's inequality (A.9) yields

$$\begin{aligned}
&\int_0^T \|\boldsymbol{u}\|_{L^4(\Omega)}^{2p}\ dt \\
&\le C\int_0^T \|\boldsymbol{u}\|_{L^4(\Omega)}^{p}\,\|\boldsymbol{u}\|_{H^{3/4}(\Omega)}^{p}\ dt \\
&\le C\int_0^T \|\boldsymbol{u}\|_{L^4(\Omega)}^{p}\,\|\boldsymbol{u}\|_{L^2(\Omega)}^{p/4}\,\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^{3p/4}\ dt \\
&\le C\,\|\boldsymbol{u}\|_{L^\infty((0,T);L^2(\Omega))}^{p/4}\int_0^T \|\boldsymbol{u}\|_{L^4(\Omega)}^{p}\,\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^{3p/4}\ dt \qquad (7.30)\\
&\le C\,\|\boldsymbol{u}\|_{L^\infty((0,T);L^2(\Omega))}^{p/4}\left(\int_0^T \|\boldsymbol{u}\|_{L^4(\Omega)}^{8p/(8-3p)}\ dt\right)^{(8-3p)/8}\left(\int_0^T \|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^{2}\ dt\right)^{3p/8}.
\end{aligned}$$

Applying the regularity assumptions of a weak solution and of Lemma 7.22 shows that all terms on the right-hand side of (7.30) are finite for $p = 2$.

Assume now that there are two weak solutions $\boldsymbol{u}', \boldsymbol{u}'' \in L^8\left(0, T; L^4(\Omega)\right)$ for the same data $\boldsymbol{f}$ and $\boldsymbol{u}_0$. Subtracting the equations for both solutions, using $\boldsymbol{u}'-\boldsymbol{u}'' \in V_{\text{div}}$ as test function, and applying (7.13) yields

$$\frac{1}{2}\frac{d}{dt}\|\boldsymbol{u}'-\boldsymbol{u}''\|_{L^2(\Omega)}^2+\nu\left\|\nabla\left(\boldsymbol{u}'-\boldsymbol{u}''\right)\right\|_{L^2(\Omega)}^2 = n\left(\boldsymbol{u}'',\boldsymbol{u}'',\boldsymbol{u}'-\boldsymbol{u}''\right)-n\left(\boldsymbol{u}',\boldsymbol{u}',\boldsymbol{u}'-\boldsymbol{u}''\right). \qquad (7.31)$$

Note that the term $(\partial_t(\boldsymbol{u}'-\boldsymbol{u}''), \boldsymbol{u}'-\boldsymbol{u}'')$ is well defined because it was shown at the beginning of the proof that $\partial_t\boldsymbol{u}', \partial_t\boldsymbol{u}'' \in L^2(0, T; V')$ The terms on the right-hand side of (7.31) will be estimated. To this end, $n(\boldsymbol{u}', \boldsymbol{u}'', \boldsymbol{u}'-\boldsymbol{u}'')$ is added and subtracted, (6.28), (6.24), (6.33) with $p = r = 4$ are used, the Sobolev imbedding $H^{3/4}(\Omega) \to L^4(\Omega)$ is used, which holds for $d = 2$ and $d = 3$, see (A.15), the interpolation theorem between Sobolev spaces (A.13) is applied, and Young's inequality (A.5) with $p = 8/7$ and $q = 8$ is used to obtain

$$\begin{aligned}
&n\left(\boldsymbol{u}'',\boldsymbol{u}'',\boldsymbol{u}'-\boldsymbol{u}''\right)-n\left(\boldsymbol{u}',\boldsymbol{u}',\boldsymbol{u}'-\boldsymbol{u}''\right)\\
&= n\left(\boldsymbol{u}''-\boldsymbol{u}',\boldsymbol{u}'',\boldsymbol{u}'-\boldsymbol{u}''\right)-n\left(\boldsymbol{u}',\boldsymbol{u}'-\boldsymbol{u}'',\boldsymbol{u}'-\boldsymbol{u}''\right)\\
&= n\left(\boldsymbol{u}''-\boldsymbol{u}',\boldsymbol{u}'',\boldsymbol{u}'-\boldsymbol{u}''\right)\\
&= -n\left(\boldsymbol{u}''-\boldsymbol{u}',\boldsymbol{u}'-\boldsymbol{u}'',\boldsymbol{u}''\right)\\
&\le \left\|\boldsymbol{u}'-\boldsymbol{u}''\right\|_{L^4(\Omega)}\left\|\nabla\left(\boldsymbol{u}'-\boldsymbol{u}''\right)\right\|_{L^2(\Omega)}\left\|\boldsymbol{u}''\right\|_{L^4(\Omega)}\\
&\le C\left\|\boldsymbol{u}'-\boldsymbol{u}''\right\|_{H^{3/4}(\Omega)}\left\|\nabla\left(\boldsymbol{u}'-\boldsymbol{u}''\right)\right\|_{L^2(\Omega)}\left\|\boldsymbol{u}''\right\|_{L^4(\Omega)}\\
&\le C\left\|\boldsymbol{u}'-\boldsymbol{u}''\right\|_{L^2(\Omega)}^{1/4}\left\|\nabla\left(\boldsymbol{u}'-\boldsymbol{u}''\right)\right\|_{L^2(\Omega)}^{7/4}\left\|\boldsymbol{u}''\right\|_{L^4(\Omega)}\\
&\le \frac{\nu}{2}\left\|\nabla\left(\boldsymbol{u}'-\boldsymbol{u}''\right)\right\|_{L^2(\Omega)}^2 + C\left\|\boldsymbol{u}''\right\|_{L^4(\Omega)}^8\left\|\boldsymbol{u}'-\boldsymbol{u}''\right\|_{L^2(\Omega)}^2. \qquad (7.32)
\end{aligned}$$

Inserting this estimate in (7.31) gives

$$\frac{d}{dt}\left\|\boldsymbol{u}'-\boldsymbol{u}''\right\|_{L^2(\Omega)}^2+\nu\left\|\nabla\left(\boldsymbol{u}'-\boldsymbol{u}''\right)\right\|_{L^2(\Omega)}^2\leq C\left\|\boldsymbol{u}''\right\|_{L^4(\Omega)}^8\left\|\boldsymbol{u}'-\boldsymbol{u}''\right\|_{L^2(\Omega)}^2. \tag{7.33}$$

The second term on the left-hand side can be estimate with zero from below. Then, applying Gronwall's lemma, Lemma A.54, yields for all times

$$\left\|\left(\boldsymbol{u}'-\boldsymbol{u}''\right)(t)\right\|_{L^2(\Omega)}^2\leq\exp\left(C\left\|\boldsymbol{u}''\right\|_{L^8(0,T;L^4(\Omega))}^8\right)\left\|\left(\boldsymbol{u}'-\boldsymbol{u}''\right)(0)\right\|_{L^2(\Omega)}^2=0,$$

since both solutions satisfy the same initial condition. Hence $\boldsymbol{u}'=\boldsymbol{u}''$ in $L^\infty(0,T;H_{\text{div}}(\Omega))$. Inserting this result in (7.33) gives for all times $\|\nabla(\boldsymbol{u}'-\boldsymbol{u}'')\|_{L^2(\Omega)}=0$ such that $\boldsymbol{u}'=\boldsymbol{u}''$ in $L^2(0,T;V_{\text{div}})$. Consequently, it is $\boldsymbol{u}'=\boldsymbol{u}''$. ■

Remark 7.23 (Generalization of Lemma 7.22) The result of Lemma 7.22 can be generalized to the statement that there is not more than one weak solution

$$\boldsymbol{u}\in L^s(0,T;L^q(\Omega))\quad\text{with}\quad s>2,q>3,\ \frac{2}{s}+\frac{3}{q}=1, \tag{7.34}$$

see Serrin (1963). A solution that satisfies (7.34) is called strong solution in the sense of Serrin.

The proof of this statement follows the lines of the proof of Lemma 7.22. First, one has to show that $\partial_t\boldsymbol{u}\in L^2(0,T;V')$. To this end, one modifies estimate (7.24), using the generalized Hölder inequality (6.33), in the form

$$\int_0^T\left(\sup_{\substack{\boldsymbol{v}\in V\\ \|\nabla\boldsymbol{v}\|_{L^2(\Omega)}=1}}|n(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v})|\right)^p dt\leq\int_0^T\int_0^T\|\boldsymbol{u}\|_{L^{2q/(q-2)}(\Omega)}^p\|\boldsymbol{u}\|_{L^q(\Omega)}^p\,dt.$$

Now, the Sobolev imbedding $H^{3/q}(\Omega)\to L^{2q/(q-2)}(\Omega)$, compare (A.15), the interpolation in Sobolev spaces (A.13), the definition of $s=(2pq)/(2q-3p)$ from (7.34), the Poincaré inequality (A.12), and Hölder's inequality (A.9) are used to obtain

$$\begin{aligned}&\int_0^T\int_0^T\|\boldsymbol{u}\|_{L^{2q/(q-2)}(\Omega)}^p\|\boldsymbol{u}\|_{L^q(\Omega)}^p\,dt\\ &\leq C\int_0^T\int_0^T\|\boldsymbol{u}\|_{H^{3/q}(\Omega)}^p\|\boldsymbol{u}\|_{L^q(\Omega)}^p\,dt\end{aligned}$$

$$
\begin{aligned}
&\le C \int_0^T \|\boldsymbol{u}\|_{L^2(\Omega)}^{2p/s} \|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^{3p/q} \|\boldsymbol{u}\|_{L^q(\Omega)}^{p} \, dt \\
&\le C \|\boldsymbol{u}\|_{L^\infty\left((0,T);L^2(\Omega)\right)}^{2p/s} \|\nabla \boldsymbol{u}\|_{L^2\left((0,T);L^2(\Omega)\right)}^{3p/q} \|\boldsymbol{u}\|_{L^s((0,T);L^q(\Omega))}^{p} < \infty,
\end{aligned}
$$

compare (7.30) for the special case $s = 8$ and $q = 4$.

For bounding the left-hand side of (7.32), principally the same tools are used. Only, the Sobolev imbedding $H^{3/q}(\Omega) \to L^{2q/(q-2)}(\Omega)$, Young's inequality (A.5) with $q = 2/(1 + 3/q)), p = 2q/(q-3)$, and the definition of s from (7.34) are applied to get

$$
\begin{aligned}
&-n\left(\boldsymbol{u}'' - \boldsymbol{u}', \boldsymbol{u}' - \boldsymbol{u}'', \boldsymbol{u}''\right) \\
&\le \left\|\boldsymbol{u}' - \boldsymbol{u}''\right\|_{L^{2q/(q-2)}(\Omega)} \left\|\nabla\left(\boldsymbol{u}' - \boldsymbol{u}''\right)\right\|_{L^2(\Omega)} \left\|\boldsymbol{u}''\right\|_{L^q(\Omega)} \\
&\le C \left\|\boldsymbol{u}' - \boldsymbol{u}''\right\|_{H^{3/q}(\Omega)} \left\|\nabla\left(\boldsymbol{u}' - \boldsymbol{u}''\right)\right\|_{L^2(\Omega)} \left\|\boldsymbol{u}''\right\|_{L^q(\Omega)} \\
&\le C \left\|\boldsymbol{u}' - \boldsymbol{u}''\right\|_{L^2(\Omega)}^{2/s} \left\|\nabla\left(\boldsymbol{u}' - \boldsymbol{u}''\right)\right\|_{L^2(\Omega)}^{1+3/q} \left\|\boldsymbol{u}''\right\|_{L^q(\Omega)} \\
&\le \frac{\nu}{2} \left\|\nabla\left(\boldsymbol{u}' - \boldsymbol{u}''\right)\right\|_{L^2(\Omega)}^2 + C \left\|\boldsymbol{u}''\right\|_{L^q(\Omega)}^s \left\|\boldsymbol{u}' - \boldsymbol{u}''\right\|_{L^2(\Omega)}^2 .
\end{aligned}
$$

With this estimate, the proof can be finished in the same way as the proof of Lemma 7.22. □

Theorem 7.24 (Uniqueness of the Weak Solution in Two Dimensions) *There is exactly one weak velocity solution of the Navier–Stokes equations in two dimensions.*

Proof The proof proceeds analogously as the proof of Lemma 7.22. The right-hand side of (7.31) has to be estimated. Bounding the right-hand side uses more or less the same tools as applied in (7.32). But in two dimensions, one can use the Sobolev imbedding $H^{1/2}(\Omega) \to L^4(\Omega)$, see (A.15), and then Young's inequality (A.5) with $p = 4/3$ and $q = 4$ to get

$$
\begin{aligned}
&n\left(\boldsymbol{u}'', \boldsymbol{u}'', \boldsymbol{u}' - \boldsymbol{u}''\right) - n\left(\boldsymbol{u}', \boldsymbol{u}', \boldsymbol{u}' - \boldsymbol{u}''\right) \\
&\le \left\|\boldsymbol{u}' - \boldsymbol{u}''\right\|_{L^4(\Omega)} \left\|\nabla\left(\boldsymbol{u}' - \boldsymbol{u}''\right)\right\|_{L^2(\Omega)} \left\|\boldsymbol{u}''\right\|_{L^4(\Omega)} \\
&\le C \left\|\boldsymbol{u}' - \boldsymbol{u}''\right\|_{H^{1/2}(\Omega)} \left\|\nabla\left(\boldsymbol{u}' - \boldsymbol{u}''\right)\right\|_{L^2(\Omega)} \left\|\boldsymbol{u}''\right\|_{L^4(\Omega)} \\
&\le C \left\|\boldsymbol{u}' - \boldsymbol{u}''\right\|_{L^2(\Omega)}^{1/2} \left\|\nabla\left(\boldsymbol{u}' - \boldsymbol{u}''\right)\right\|_{L^2(\Omega)}^{3/2} \left\|\boldsymbol{u}''\right\|_{L^4(\Omega)} \\
&\le \frac{\nu}{2} \left\|\nabla\left(\boldsymbol{u}' - \boldsymbol{u}''\right)\right\|_{L^2(\Omega)}^2 + C \left\|\boldsymbol{u}''\right\|_{L^4(\Omega)}^4 \left\|\boldsymbol{u}' - \boldsymbol{u}''\right\|_{L^2(\Omega)}^2 .
\end{aligned}
$$

Since by (7.25), it is known that in two dimensions $\boldsymbol{u}'' \in L^4\left((0,T); L^4(\Omega)\right)$, the proof can be finished in the same way as the proof of Lemma 7.22. ■

Remark 7.25 (On the Uniqueness of the Weak Solution in Three Dimensions) The uniqueness of a weak solution in the three-dimensional case is an open problem. Only the uniqueness of a strong solution in the sense of Serrin is known, compare Remark 7.23. However, the existence of a strong solution cannot be proved, only the existence of a weak solution. These open questions are strongly linked to one of the major mathematical challenges that were formulated in one of the so-called *Millennium Problems* of the Clay Mathematics Institute, see Fefferman (2000).

From the mathematical point of view, the difference of the two- and three-dimensional case arises from the dependency of Sobolev imbeddings on the dimension of the domain, compare the proofs of Lemma 7.22 and Theorem 7.24.

□

Remark 7.26 (Local Existence and Uniqueness Results for Small Data) In three dimensions, one can prove that there exists a unique solution of the Navier–Stokes equations with higher regularity and in a time interval $(0, T')$, $T' > 0$, if the data $\boldsymbol{u}_0$ and $\boldsymbol{f}$ are sufficiently small, e.g., see Sohr (2001, Chap. V, Theorem 4.2.2). The time T' depends on the size of the data. This solution is sometimes called local solution.

□

Remark 7.27 (The Pressure) Given a weak velocity solution $\boldsymbol{u}$ of the Navier–Stokes equations, a pressure p can be defined such that $(\boldsymbol{u}, p)$ satisfy a weak form of the Navier–Stokes equations. One can find in the literature several results concerning the regularity of the pressure, e.g., Galdi (2000, Sect. 2), Sohr (2001, Sect. V.1.7), or Temam (1984, p. 307). □

Remark 7.28 (An Equation for the Pressure for $t > 0$) A possible way to construct an equation for the pressure, which is the most common one, starts by assuming that all functions in the Navier–Stokes equations (7.1) are sufficiently smooth. Then, the divergence operator is applied to the momentum equation. Using that $\boldsymbol{u}$ is divergence-free, such that

$$\nabla \cdot \partial_t \boldsymbol{u} = \partial_t (\nabla \cdot \boldsymbol{u}) = 0$$

and $\nabla \cdot \Delta \boldsymbol{u} = 0$, compare (4.123), gives

$$-\nabla \cdot \nabla p = -\Delta p = -\nabla \cdot \boldsymbol{f} + \nabla \cdot ((\boldsymbol{u} \cdot \nabla))\boldsymbol{u}) \,. \tag{7.35}$$

This equation is a Poisson equation for the pressure, a so-called pressure Poisson equation. To get a well-posed problem, (7.35) has to be equipped with appropriate boundary conditions. To this end, one considers the restriction of the momentum equation in (7.1) to the boundary Γ, multiplies this restriction with the outward pointing unit normal vector $\boldsymbol{n}$, and uses that $\boldsymbol{u}$ vanishes at the boundary for $t > 0$, which yields

$$\nabla p \cdot \boldsymbol{n} = (\boldsymbol{f} + \nu \Delta \boldsymbol{u}) \cdot \boldsymbol{n} \quad \text{on } \Gamma. \tag{7.36}$$

Finally, the pressure is contained in $L_0^2(\Omega)$ such that it has to satisfy

$$\int_\Omega p(\boldsymbol{x})\, d\boldsymbol{x} = 0. \tag{7.37}$$

Problem (7.35)–(7.37) admits a unique solution. □

Remark 7.29 (The Pressure at the Initial Time) A pressure at the initial time is not part of the definition of the problem. For obtaining an equation, one uses in principal the same way as described in Remark 7.28, however, the derivation of the boundary condition has to take into account that generally $\boldsymbol{u}_0$ does not vanish on Γ. By Definition 7.6, it is $\boldsymbol{u}_0 \in H_{\text{div}}(\Omega)$ such that only the normal component of $\boldsymbol{u}_0$ vanishes on Γ. Since the normal component of the velocity vanishes also for all times $t > 0$, one finds that $\partial_t \boldsymbol{u}_0 \cdot \boldsymbol{n}|_\Gamma = 0$. Hence, one gets the problem

$$\begin{aligned} -\Delta p_0 &= -\nabla \cdot \boldsymbol{f}(0,\cdot) + \nabla \cdot ((\boldsymbol{u}_0 \cdot \nabla)\, \boldsymbol{u}_0) \quad \text{in } \Omega, \\ \nabla p_0 \cdot \boldsymbol{n} &= (\boldsymbol{f}(0,\cdot) + \nu \Delta \boldsymbol{u}_0 - (\boldsymbol{u}_0 \cdot \nabla)\, \boldsymbol{u}_0) \cdot \boldsymbol{n} \quad \text{on } \Gamma, \\ \int_\Omega p_0(\boldsymbol{x})\, d\boldsymbol{x} &= 0. \end{aligned} \tag{7.38}$$

It was shown in Heywood and Rannacher (1982) that this problem is well-posed under certain regularity assumptions on $\boldsymbol{f}$ and $\boldsymbol{u}_0$ and it defines a unique pressure at the initial time. □

Remark 7.30 (Regularity for $t \to +0$, Compatibility Conditions) Restricting the momentum equation for $t = 0$ to the boundary, then one finds for the initial pressure computed with (7.38) that

$$\nabla p_0|_\Gamma = (\boldsymbol{f}(0,\cdot) - \partial_t \boldsymbol{u}_0 + \nu \Delta \boldsymbol{u}_0 - (\boldsymbol{u}_0 \cdot \nabla)\, \boldsymbol{u}_0)|_\Gamma .$$

If one requires that there is a continuous transition for the velocity as $t \to +0$, then $\partial_t \boldsymbol{u}_0|_\Gamma = \boldsymbol{0}$ since $\boldsymbol{u}(t)|_\Gamma = 0$ for all $t > 0$. In this case, the initial pressure has to satisfy the boundary condition

$$\nabla p_0|_\Gamma = (\boldsymbol{f}(0,\cdot) + \nu \Delta \boldsymbol{u}_0 - (\boldsymbol{u}_0 \cdot \nabla)\, \boldsymbol{u}_0)|_\Gamma . \tag{7.39}$$

However, the pressure Poisson equation from (7.38) together with the boundary condition (7.39) constitute an over-determined problem. Only the normal component of the gradient of the pressure can be prescribed at the boundary but not the whole gradient. Note that the condition on the whole gradient can be split in a condition on the normal component, as in (7.38), and a condition on each of the linearly independent tangential vectors, in the same way as described in Remark 2.25. Usually, this over-determined problem does not possess a solution. Hence, it is generally

$$\lim_{t\to+0} \partial_t \boldsymbol{u}|_\Gamma = \lim_{t\to+0} (\boldsymbol{f} + \nu \Delta \boldsymbol{u} - (\boldsymbol{u} \cdot \nabla)\, \boldsymbol{u} - \nabla p)|_\Gamma \neq \boldsymbol{0}.$$

The over-determined problem is solvable only if certain compatibility conditions of $\boldsymbol{f}$ and $\boldsymbol{u}_0$ are satisfied, see Temam (1982) for details. The difficulty is not that generally $\boldsymbol{u}_0$ does not vanish on Γ. Even for $\boldsymbol{u}_0$ with compact support in Ω, where (7.39) reduces to $\nabla p_0|_\Gamma = \boldsymbol{f}(0,\cdot)|_\Gamma$, such compatibility conditions are required. These compatibility conditions are nonlocal, since they include the right-hand side of the pressure Poisson equation, and they are virtually uncheckable for given data, see Heywood and Rannacher (1982). They are rarely satisfied, even in modeling simple problems, compare Heywood and Rannacher (1988).

The reduced regularity at the initial time can be quantified, e.g., it was proved in Heywood and Rannacher (1982, Theorem 2.5) for $k \ge 2$

$$\|\boldsymbol{u}(t)\|_{H^k(\Omega)} + \|\partial_t \boldsymbol{u}(t)\|_{H^{k-2}(\Omega)} \le C\left(\min\{t,1\}\right)^{1-k/2}, \quad t > 0.$$

Thus, higher order norms of $\boldsymbol{u}$ and $\partial_t \boldsymbol{u}$ blow up as $t \to +0$. □

7.2 Finite Element Error Analysis: The Time-Continuous Case

Remark 7.31 (Motivation) The error analysis in the time-continuous case does not consider a discretization of the temporal derivative in the momentum equation in (7.1). Thus, one concentrates on the errors that are introduced by the spatial discretization and the dependency of the constants in the error bounds on the viscosity ν. □

Remark 7.32 (The Weak Formulation) In the finite element error analysis, a weak formulation of the time-dependent Navier–Stokes equations is considered that does not apply an integration by parts with respect to the temporal derivative. Consider (7.1) with the boundary condition (7.2). The weak formulation reads as follows: Find $\boldsymbol{u} : (0,T] \to V$ and $p : (0,T] \to Q$ such that

$$(\partial_t \boldsymbol{u}, \boldsymbol{v}) + (\nu \nabla \boldsymbol{u}, \nabla \boldsymbol{v}) + n\left(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}\right) - (\nabla \cdot \boldsymbol{v}, p) + (\nabla \cdot \boldsymbol{u}, q) = \langle \boldsymbol{f}, \boldsymbol{v}\rangle_{V',V} \qquad (7.40)$$

for all $(\boldsymbol{v}, q) \in V \times Q$ and $\boldsymbol{u}(0,\boldsymbol{x}) = \boldsymbol{u}_0(\boldsymbol{x}) \in H_{\mathrm{div}}(\Omega)$. In (7.40), the nonlinear term $n\left(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}\right)$ can be chosen as any of the terms that are introduced in Remark 6.8.

By definition, the regularity (7.5) is known. The satisfaction of the energy inequality (7.27) and the stability estimate (7.28) were proved in Lemma 7.21. For the finite element error analysis, it will turn out that the regularity (7.5) is not sufficient. Higher regularity assumptions are used, see Theorem 7.35. □

Remark 7.33 (Time-Continuous Galerkin Finite Element Formulation) This section considers inf-sup stable and conforming finite element spaces. Let $V^h \subset V$ and $Q^h \subset Q$ and let V^h and Q^h satisfy the discrete inf-sup condition (3.51), then the time-continuous Galerkin finite element formulation reads as follows: Find $\boldsymbol{u}^h$:

$(0, T] \to V^h$ and $p^h : (0, T] \to Q^h$ such that

$$\begin{aligned}\left(\partial_t \boldsymbol{u}^h, \boldsymbol{v}^h\right) + (\nu \nabla \boldsymbol{u}^h, \nabla \boldsymbol{v}^h) + n\left(\boldsymbol{u}^h, \boldsymbol{u}^h, \boldsymbol{v}^h\right) \\ - \left(\nabla \cdot \boldsymbol{v}^h, p^h\right) + \left(\nabla \cdot \boldsymbol{u}^h, q^h\right) = \langle \boldsymbol{f}, \boldsymbol{v}^h \rangle_{V',V} \qquad (7.41)\end{aligned}$$

for all $(\boldsymbol{v}^h, q^h) \in V^h \times Q^h$ and $\boldsymbol{u}^h(0, \boldsymbol{x}) = \boldsymbol{u}_0^h(\boldsymbol{x}) \in V^h$, where $\boldsymbol{u}_0^h(\boldsymbol{x})$ is an approximation of $\boldsymbol{u}_0(\boldsymbol{x})$, for instance an appropriate interpolation (if the initial condition is sufficiently smooth) or a projection. For the domain, the usual assumptions have to be made: Ω should be a bounded domain with polyhedral and Lipschitz continuous boundary.

Like in the stability analysis of the steady-state equations, it is important that the nonlinear term vanishes if its second and third argument are identical, see Lemma 6.10. Thus, the nonlinear term for the stability estimates of (7.41) might be $n_{\text{skew}}\left(\boldsymbol{u}^h, \boldsymbol{u}^h, \boldsymbol{v}^h\right)$, $n_{\text{rot}}\left(\boldsymbol{u}^h, \boldsymbol{u}^h, \boldsymbol{v}^h\right)$, or $n_{\text{div}}\left(\boldsymbol{u}^h, \boldsymbol{u}^h, \boldsymbol{v}^h\right)$.

In contrast to the finite element error analysis of the steady-state Navier–Stokes equations, the $L^2(\Omega)$ norm of the velocity plays an important role in the error analysis of the time-dependent problem. This fact is reflected by the use of an estimate of the convective term that involves the $L^2(\Omega)$ norm of the first argument, see Lemma 6.14. This estimate is only true for the convective and the skew-symmetric form. Since the results of the stability analysis are applied in the finite element error analysis, this analysis will be carried out for $n_{\text{skew}}(\cdot, \cdot, \cdot)$ only.

In practical simulations, $n_{\text{conv}}(\cdot, \cdot, \cdot)$ is often used. □

Lemma 7.34 (Existence, Uniqueness, and Stability of the Finite Element Solution) *Let $\boldsymbol{u}_0^h \in V_{\text{div}}^h$ and $\boldsymbol{f} \in L^2(0, t; V')$, then the finite element problem (7.41) has a unique solution $\left(\boldsymbol{u}^h, p^h\right) \in V^h \times Q^h$, where V^h and Q^h are assumed to satisfy the discrete inf-sup condition (3.51). If the nonlinear terms $n_{\text{skew}}\left(\boldsymbol{u}^h, \boldsymbol{u}^h, \boldsymbol{v}^h\right)$, $n_{\text{rot}}\left(\boldsymbol{u}^h, \boldsymbol{u}^h, \boldsymbol{v}^h\right)$, or $n_{\text{div}}\left(\boldsymbol{u}^h, \boldsymbol{u}^h, \boldsymbol{v}^h\right)$ are used, then it holds for all $t \in (0, T]$ that*

$$\left\|\boldsymbol{u}^h(t)\right\|_{L^2(\Omega)}^2 + \nu \left\|\nabla \boldsymbol{u}^h\right\|_{L^2(0,t;L^2(\Omega))}^2 \le \left\|\boldsymbol{u}_0^h\right\|_{L^2(\Omega)}^2 + \frac{1}{\nu} \|\boldsymbol{f}\|_{L^2(0,t;V')}^2 . \qquad (7.42)$$

Proof Consider first the Galerkin finite element method in V_{div}^h: Find $\boldsymbol{u}^h \in V_{\text{div}}^h$ such that

$$\left(\partial_t \boldsymbol{u}^h, \boldsymbol{v}^h\right) + (\nu \nabla \boldsymbol{u}^h, \nabla \boldsymbol{v}^h) + n\left(\boldsymbol{u}^h, \boldsymbol{u}^h, \boldsymbol{v}^h\right) = \langle \boldsymbol{f}, \boldsymbol{v}^h \rangle_{V',V} \quad \forall\ \boldsymbol{v}^h \in V_{\text{div}}^h.$$

This problem is a problem in a finite-dimensional space, exactly as problem (7.6) or equivalently (7.7)–(7.9). Now, the existence and uniqueness of a velocity solution can be proved with the same arguments as in the proof of Lemma 7.12. The existence and uniqueness of a corresponding pressure follows from the discrete inf-sup condition (3.51).

Estimate (7.42) is the analog of estimate (7.11). Note that in the derivation of (7.11) the skew-symmetry of the nonlinear term was used, which is given for finite element spaces usually only for the forms of the convective term mentioned in the formulation of the lemma. In contrast to (7.11), the estimate of the projection for the initial velocity is not yet applied in (7.42). ■

Theorem 7.35 (Finite Element Error Estimate for the Velocity of the Continuous-in-Time Galerkin Finite Element Method) *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, be a bounded domain with polyhedral and Lipschitz continuous boundary, let $\boldsymbol{f} \in L^2(0,T;V')$, $\boldsymbol{u}_0 \in H_{\mathrm{div}}(\Omega)$, and $\boldsymbol{u}_0^h \in V_{\mathrm{div}}^h$. In addition, the regularities*

$$\partial_t \boldsymbol{u} \in L^2\left(0,T;V'\right), \quad \nabla \boldsymbol{u} \in L^4\left(0,T;L^2(\Omega)\right), \quad p \in L^2\left(0,T;L^2(\Omega)\right) \tag{7.43}$$

of the solution of (7.40) are assumed. Let $\left(\boldsymbol{u}^h, p^h\right) \in V^h \times Q^h$ be the solution of (7.41), where V^h and Q^h satisfy the discrete inf-sup condition (3.51) and the skew-symmetric nonlinear term is used, then the following error estimate holds for all $t \in (0,T]$

$$\begin{aligned}
&\left\|\left(\boldsymbol{u}-\boldsymbol{u}^h\right)(t)\right\|_{L^2(\Omega)}^2 + \nu\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2\left(0,t;L^2(\Omega)\right)}^2 \\
&\le C\Bigg\{\left\|\left(\boldsymbol{u}-I_{\mathrm{St}}^h\boldsymbol{u}\right)(t)\right\|_{L^2(\Omega)}^2 + \nu\left\|\nabla\left(\boldsymbol{u}-I_{\mathrm{St}}^h\boldsymbol{u}\right)\right\|_{L^2\left(0,t;L^2(\Omega)\right)}^2 \\
&\quad + \exp\left(\frac{C}{\nu^3}\left\|\nabla\boldsymbol{u}\right\|_{L^4\left(0,t;L^2(\Omega)\right)}^4\right)\Bigg[\left\|\boldsymbol{u}_0^h - I_{\mathrm{St}}^h\boldsymbol{u}(0)\right\|_{L^2(\Omega)}^2 \\
&\quad + \frac{1}{\nu}\Bigg(\left\|\partial_t\left(\boldsymbol{u}-I_{\mathrm{St}}^h\boldsymbol{u}\right)\right\|_{L^2(0,t;V')}^2 + \left\|\nabla\left(\boldsymbol{u}-I_{\mathrm{St}}^h\boldsymbol{u}\right)\right\|_{L^4\left(0,t;L^2(\Omega)\right)}^2\left\|\nabla\boldsymbol{u}\right\|_{L^4\left(0,t;L^2(\Omega)\right)}^2 \\
&\quad + \inf_{q^h\in L^2(0,t;Q^h)}\left\|p-q^h\right\|_{L^2\left(0,t;L^2(\Omega)\right)}^2\Bigg) \\
&\quad + \frac{1}{\nu^{3/2}}\left(\left\|\boldsymbol{u}_0^h\right\|_{L^2(\Omega)}^2 + \frac{1}{\nu}\left\|\boldsymbol{f}\right\|_{L^2(0,t;V')}^2\right)\left\|\nabla\left(\boldsymbol{u}-I_{\mathrm{St}}^h\boldsymbol{u}\right)\right\|_{L^4\left(0,t;L^2(\Omega)\right)}^2\Bigg]\Bigg\},
\end{aligned} \tag{7.44}$$

where $I_{\mathrm{St}}^h\boldsymbol{u}(t)$ is the Stokes projection at time t with $p=0$, see (4.54), for which

$$\partial_t I_{\mathrm{St}}^h\boldsymbol{u} \in L^2\left(0,T;V'\right) \tag{7.45}$$

is assumed.

Proof The proof proceeds in several steps:

1. derivation of an error equation and splitting of the error,
2. estimate all terms on the right-hand side of the error equation,
3. application of Gronwall's lemma,
4. application of the triangle inequality.

1. *Derivation of an error equation and splitting of the error.* An error equation is derived by subtracting the weak form and the finite element formulation for test functions that can be applied in both equations. As usual, the error is decomposed into an interpolation error and a finite element remainder

$$\boldsymbol{e}(t) = \boldsymbol{u}(t) - \boldsymbol{u}^h(t) = \left(\boldsymbol{u}(t) - I_{\mathrm{St}}^h \boldsymbol{u}(t)\right) + \left(I_{\mathrm{St}}^h \boldsymbol{u}(t) - \boldsymbol{u}^h(t)\right) = \boldsymbol{\eta}(t) - \boldsymbol{\phi}^h(t). \tag{7.46}$$

From the estimate (4.55) and the regularity assumptions (7.43), it follows that

$$\nabla I_{\mathrm{St}}^h \boldsymbol{u} \in L^4\left(0, T; L^2(\Omega)\right). \tag{7.47}$$

For simplicity of notation, the argument will be neglected in the following. Now, (7.40) and (7.41) are considered for test functions $\boldsymbol{v}^h \in V_{\mathrm{div}}^h \subset V^h \subset V$ and $q^h \in Q^h \subset Q$. Subtracting both equations and using that $\boldsymbol{u}^h$ is discretely divergence-free leads to

$$\left(\partial_t \boldsymbol{e}, \boldsymbol{v}^h\right) + \left(\nu \nabla \boldsymbol{e}, \nabla \boldsymbol{v}^h\right) + n_{\mathrm{skew}}\left(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}^h\right) - n_{\mathrm{skew}}\left(\boldsymbol{u}^h, \boldsymbol{u}^h, \boldsymbol{v}^h\right) - \left(\nabla \cdot \boldsymbol{v}^h, p\right) = 0$$

for all $\boldsymbol{v}^h \in V_{\mathrm{div}}^h$. Since $\boldsymbol{v}^h$ is discretely divergence-free, it is $\left(\nabla \cdot \boldsymbol{v}^h, q^h\right) = 0$ for all $q^h \in Q^h$ and this term can be added to this equation. Rearranging terms yields

$$\begin{aligned} &\left(\partial_t \boldsymbol{\phi}^h, \boldsymbol{v}^h\right) + \left(\nu \nabla \boldsymbol{\phi}^h, \nabla \boldsymbol{v}^h\right) \\ &= \left(\partial_t \boldsymbol{\eta}, \boldsymbol{v}^h\right) + \left(\nu \nabla \boldsymbol{\eta}, \nabla \boldsymbol{v}^h\right) + n_{\mathrm{skew}}\left(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}^h\right) - n_{\mathrm{skew}}\left(\boldsymbol{u}^h, \boldsymbol{u}^h, \boldsymbol{v}^h\right) - \left(\nabla \cdot \boldsymbol{v}^h, p - q^h\right) \end{aligned} \tag{7.48}$$

for all $\left(\boldsymbol{v}^h, q^h\right) \in V_{\mathrm{div}}^h \times Q^h$. Note that by the definition of the Stokes projection and by $\boldsymbol{v}^h \in V_{\mathrm{div}}^h$, it is $\left(\nu \nabla \boldsymbol{\eta}, \nabla \boldsymbol{v}^h\right) = 0$. Since $I_{\mathrm{St}}^h \boldsymbol{u}$ is as Stokes projection discretely divergence-free and $\boldsymbol{u}^h$ is as solution of (7.41) discretely divergence-free, it follows that $\boldsymbol{\phi}^h$ is also discretely divergence-free. Hence, one can choose $\boldsymbol{v}^h = \boldsymbol{\phi}^h$ and obtains the error equation

$$\begin{aligned} &\frac{1}{2} \frac{d}{dt} \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 + \nu \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 \\ &= \left(\partial_t \boldsymbol{\eta}, \boldsymbol{\phi}^h\right) + n_{\mathrm{skew}}\left(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{\phi}^h\right) - n_{\mathrm{skew}}\left(\boldsymbol{u}^h, \boldsymbol{u}^h, \boldsymbol{\phi}^h\right) - \left(\nabla \cdot \boldsymbol{\phi}^h, p - q^h\right) \end{aligned} \tag{7.49}$$

for all $q^h \in Q^h$.

2. *Estimate all terms on the right-hand side of the error equation* (7.49). The goal of these estimates consists in absorbing in the left-hand side of (7.49) as many terms as possible that contain $\boldsymbol{\phi}^h$ on the right-hand side. There is not much choice for doing this job, in principle only the second term on the left-hand side can be used for this purpose.

The linear terms in (7.49) are estimated with the usual tools: dual pairing or Cauchy–Schwarz inequality (A.10), and Young's inequality (A.5). One obtains

$$\left|\left(\partial_t \boldsymbol{\eta}, \boldsymbol{\phi}^h\right)\right| \leq \|\partial_t \boldsymbol{\eta}\|_{V'} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \leq \frac{2}{\nu} \|\partial_t \boldsymbol{\eta}\|_{V'}^2 + \frac{\nu}{8} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2$$

and, using in addition (3.41),

$$\begin{aligned}\left|\left(\nabla \cdot \boldsymbol{\phi}^h, p - q^h\right)\right| &\leq \left\|\nabla \cdot \boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \left\|p - q^h\right\|_{L^2(\Omega)} \\ &\leq \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \left\|p - q^h\right\|_{L^2(\Omega)} \\ &\leq \frac{2}{\nu} \left\|p - q^h\right\|_{L^2(\Omega)}^2 + \frac{\nu}{8} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 .\end{aligned}$$

The nonlinear terms are decomposed as in (6.65). One obtains

$$\begin{aligned}&\left|n_{\text{skew}}\left(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\boldsymbol{u}^h, \boldsymbol{u}^h, \boldsymbol{\phi}^h\right)\right| \\ &\leq \left|n_{\text{skew}}\left(\boldsymbol{\eta}, \boldsymbol{u}, \boldsymbol{\phi}^h\right)\right| + \left|n_{\text{skew}}\left(\boldsymbol{\phi}^h, \boldsymbol{u}, \boldsymbol{\phi}^h\right)\right| + \left|n_{\text{skew}}\left(\boldsymbol{u}^h, \boldsymbol{\eta}, \boldsymbol{\phi}^h\right)\right| .\end{aligned}$$

Applying (6.41) with $s = 1/2$ and Young's inequality gives

$$\begin{aligned}\left|n_{\text{skew}}\left(\boldsymbol{\eta}, \boldsymbol{u}, \boldsymbol{\phi}^h\right)\right| &\leq C \|\boldsymbol{\eta}\|_{L^2(\Omega)}^{1/2} \|\nabla \boldsymbol{\eta}\|_{L^2(\Omega)}^{1/2} \|\nabla \boldsymbol{u}\|_{L^2(\Omega)} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \\ &\leq \frac{C}{\nu} \|\boldsymbol{\eta}\|_{L^2(\Omega)} \|\nabla \boldsymbol{\eta}\|_{L^2(\Omega)} \|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^2 + \frac{\nu}{8} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 ,\end{aligned} \tag{7.50}$$

$$\begin{aligned}\left|n_{\text{skew}}\left(\boldsymbol{\phi}^h, \boldsymbol{u}, \boldsymbol{\phi}^h\right)\right| &\leq \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^{1/2} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^{1/2} \|\nabla \boldsymbol{u}\|_{L^2(\Omega)} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \\ &\leq \frac{C}{\nu^3} \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 \|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^4 + \frac{\nu}{8} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 ,\end{aligned} \tag{7.51}$$

and

$$\begin{aligned}\left|n_{\text{skew}}\left(\boldsymbol{u}^h, \boldsymbol{\eta}, \boldsymbol{\phi}^h\right)\right| &\leq C \left\|\boldsymbol{u}^h\right\|_{L^2(\Omega)}^{1/2} \left\|\nabla \boldsymbol{u}^h\right\|_{L^2(\Omega)}^{1/2} \|\nabla \boldsymbol{\eta}\|_{L^2(\Omega)} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \\ &\leq \frac{C}{\nu} \left\|\boldsymbol{u}^h\right\|_{L^2(\Omega)} \left\|\nabla \boldsymbol{u}^h\right\|_{L^2(\Omega)} \|\nabla \boldsymbol{\eta}\|_{L^2(\Omega)}^2 + \frac{\nu}{8} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 .\end{aligned} \tag{7.52}$$

Inserting all estimates in (7.49) yields

$$\begin{aligned}&\frac{1}{2} \frac{d}{dt} \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 + \frac{3\nu}{8} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 \\ &\leq \frac{2}{\nu} \|\partial_t \boldsymbol{\eta}\|_{V'}^2 + \frac{2}{\nu} \left\|p - q^h\right\|_{L^2(\Omega)}^2\end{aligned}$$

$$
+\frac{C}{\nu}\left(\|\boldsymbol{\eta}\|_{L^2(\Omega)}\|\nabla\boldsymbol{\eta}\|_{L^2(\Omega)}\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^2+\left\|\boldsymbol{u}^h\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{u}^h\right\|_{L^2(\Omega)}\|\nabla\boldsymbol{\eta}\|_{L^2(\Omega)}^2\right)
+\frac{C}{\nu^3}\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^4\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2
$$

for all $q^h \in Q^h$.

3. *Application of Gronwall's Lemma A.54*. Next, this estimate is integrated in $(0,t)$ leading to

$$
\begin{aligned}
&\left\|\boldsymbol{\phi}^h(t)\right\|_{L^2(\Omega)}^2+\frac{3\nu}{4}\left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(0,t;L^2(\Omega))}^2\\
&\le\left\|\boldsymbol{\phi}^h(0)\right\|_{L^2(\Omega)}^2+\frac{2}{\nu}\|\partial_t\boldsymbol{\eta}\|_{L^2(0,t;V')}^2+\frac{2}{\nu}\left\|p-q^h\right\|_{L^2(0,t;L^2(\Omega))}^2\\
&\quad+\frac{C}{\nu}\int_0^t\|\boldsymbol{\eta}\|_{L^2(\Omega)}\|\nabla\boldsymbol{\eta}\|_{L^2(\Omega)}\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^2\,d\tau\\
&\quad+\frac{C}{\nu}\int_0^t\left\|\boldsymbol{u}^h\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{u}^h\right\|_{L^2(\Omega)}\|\nabla\boldsymbol{\eta}\|_{L^2(\Omega)}^2\,d\tau\\
&\quad+\frac{C}{\nu^3}\int_0^t\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^4\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2\,d\tau
\end{aligned}
\tag{7.53}
$$

for all $q^h \in Q^h$. Before Gronwall's lemma can be applied to estimate (7.53), it has to be checked if the assumptions for its application are satisfied. All terms on the right-hand side of (7.53) are non-negative. From the regularity assumptions (7.43) and (7.45), it follows immediately that the first two terms on the right-hand side of (7.53) are well defined. For the third term, it follows with Poincaré's inequality (A.12) and the Cauchy–Schwarz inequality that

$$
\begin{aligned}
&\int_0^t\|\boldsymbol{\eta}\|_{L^2(\Omega)}\|\nabla\boldsymbol{\eta}\|_{L^2(\Omega)}\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^2\,d\tau\\
&\le C\int_0^t\|\nabla\boldsymbol{\eta}\|_{L^2(\Omega)}^2\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^2\,d\tau\\
&\le C\int_0^t\left(\|\nabla\boldsymbol{\eta}\|_{L^2(\Omega)}^4\,d\tau\right)^{1/2}\left(\int_0^t\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^4\,d\tau\right)^{1/2}\\
&=C\|\nabla\boldsymbol{\eta}\|_{L^4(0,t;L^2(\Omega))}^2\|\nabla\boldsymbol{u}\|_{L^4(0,t;L^2(\Omega))}^2<\infty,
\end{aligned}
\tag{7.54}
$$

because of the regularity assumptions (7.43) and of (7.47). For the fourth term on the right-hand side of (7.53), one obtains with the stability estimate (7.42) and the Cauchy–Schwarz inequality

$$\begin{aligned}
&\int_0^t \left\| \boldsymbol{u}^h \right\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{u}^h \right\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{\eta} \right\|_{L^2(\Omega)}^2 \, d\tau \\
&\le \left\| \boldsymbol{u}^h \right\|_{L^\infty(0,t;L^2(\Omega))} \int_0^t \left\| \nabla \boldsymbol{u}^h \right\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{\eta} \right\|_{L^2(\Omega)}^2 \, d\tau \\
&\le \left\| \boldsymbol{u}^h \right\|_{L^\infty(0,t;L^2(\Omega))} \left\| \nabla \boldsymbol{u}^h \right\|_{L^2(0,t;L^2(\Omega))} \left\| \nabla \boldsymbol{\eta} \right\|_{L^4(0,t;L^2(\Omega))}^2 \\
&\le \frac{C}{\nu^{1/2}} \left(\left\| \boldsymbol{u}_0^h \right\|_{L^2(\Omega)}^2 + \frac{1}{\nu} \left\| \boldsymbol{f} \right\|_{L^2(0,t;V')}^2 \right) \left\| \nabla \boldsymbol{\eta} \right\|_{L^4(0,t;L^2(\Omega))}^2 < \infty \qquad (7.55)
\end{aligned}$$

as a consequence of the regularity assumptions (7.43) and of (7.47). From (7.43), it follows that

$$\int_0^t \left\| \nabla \boldsymbol{u} \right\|_{L^2(\Omega)}^4 \, d\tau < \infty.$$

Hence, all conditions for the application of Gronwall's lemma are satisfied and one obtains for all $t \in (0, T]$ and all $q^h \in Q^h$

$$\begin{aligned}
&\left\| \boldsymbol{\phi}^h(t) \right\|_{L^2(\Omega)}^2 + \nu \left\| \nabla \boldsymbol{\phi}^h \right\|_{L^2(0,t;L^2(\Omega))}^2 \\
&\le C \exp\left(\frac{C}{\nu^3} \left\| \nabla \boldsymbol{u} \right\|_{L^4(0,t;L^2(\Omega))}^4 \right) \qquad (7.56) \\
&\quad \times \Bigg[\left\| \boldsymbol{\phi}^h(0) \right\|_{L^2(\Omega)}^2 + \nu^{-1} \Big(\left\| \partial_t \boldsymbol{\eta} \right\|_{L^2(0,t;V')}^2 + \left\| \nabla \boldsymbol{\eta} \right\|_{L^4(0,t;L^2(\Omega))}^2 \left\| \nabla \boldsymbol{u} \right\|_{L^4(0,t;L^2(\Omega))}^2 \\
&\quad + \left\| p - q^h \right\|_{L^2(0,t;L^2(\Omega))}^2 \Big) \\
&\quad + \nu^{-3/2} \left(\left\| \boldsymbol{u}_0^h \right\|_{L^2(\Omega)}^2 + \frac{1}{\nu} \left\| \boldsymbol{f} \right\|_{L^2(0,t;V')}^2 \right) \left\| \nabla \boldsymbol{\eta} \right\|_{L^4(0,t;L^2(\Omega))}^2 \Bigg].
\end{aligned}$$

4. *Application of the triangle inequality.* The triangle inequality gives

$$\begin{aligned}
&\left\| \boldsymbol{e}(t) \right\|_{L^2(\Omega)}^2 + \nu \left\| \nabla \boldsymbol{e} \right\|_{L^2(0,t;L^2(\Omega))}^2 \\
&\le 2 \left(\left\| \boldsymbol{\eta}(t) \right\|_{L^2(\Omega)}^2 + \nu \left\| \nabla \boldsymbol{\eta} \right\|_{L^2(0,t;L^2(\Omega))}^2 + \left\| \boldsymbol{\phi}^h(t) \right\|_{L^2(\Omega)}^2 + \nu \left\| \nabla \boldsymbol{\phi}^h \right\|_{L^2(0,t;L^2(\Omega))}^2 \right).
\end{aligned}$$

Inserting (7.56) concludes the proof. ■

Remark 7.36 (On Error Estimate (7.44)) Estimate (7.44) has the typical form of a finite element error estimate. On the right-hand side one finds interpolation errors, data of the problem, and norms of the solution of the continuous problem.

- The term

$$\left\| \boldsymbol{u}_0^h - I_{\mathrm{St}}^h \boldsymbol{u}(0) \right\|_{L^2(\Omega)} \le \left\| \boldsymbol{u}_0 - I_{\mathrm{St}}^h \boldsymbol{u}(0) \right\|_{L^2(\Omega)} + \left\| \boldsymbol{u}_0 - \boldsymbol{u}_0^h \right\|_{L^2(\Omega)}$$

  states that the initial condition has to be approximated sufficiently accurately.
- From Remark 4.44, it follows that

$$\left\| \partial_t \left(\boldsymbol{u} - I_{\mathrm{St}}^h \boldsymbol{u} \right) \right\|_{L^2(0,t;V')} = \left\| \partial_t \boldsymbol{u} - I_{\mathrm{St}}^h \left(\partial_t \boldsymbol{u} \right) \right\|_{L^2(0,t;V')} .$$

- The dominating interpolation errors are

$$\left\| \nabla \left(\boldsymbol{u} - I_{\mathrm{St}}^h \boldsymbol{u} \right) \right\|_{L^2\left(0,t;L^2(\Omega)\right)}, \quad \left\| \nabla \left(\boldsymbol{u} - I_{\mathrm{St}}^h \boldsymbol{u} \right) \right\|_{L^4\left(0,t;L^2(\Omega)\right)},$$
$$\inf_{q^h \in L^2(0,t;Q^h)} \left\| p - q^h \right\|_{L^2\left(0,t;L^2(\Omega)\right)},$$

  where the errors for the Stokes projection are estimated in (4.57) (with the pressure term vanishing since $p = 0$).

□

Remark 7.37 (On the Regularity Assumptions of Theorem 7.35) The regularity assumptions (7.43) imply the uniqueness of the weak solution. Applying the Sobolev imbedding $H^1(\Omega) \to L^6(\Omega)$, see (A.22), and Poincaré's inequality (A.12) yields

$$\int_0^T \|\boldsymbol{u}\|_{L^6(\Omega)}^4 \, d\tau \le C \int_0^T \|\boldsymbol{u}\|_{H^1(\Omega)}^4 \, d\tau \le C \int_0^T \|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^4 \, d\tau < \infty,$$

such that $\boldsymbol{u} \in L^4\left(0, T; L^6(\Omega)\right)$. Hence, condition (7.34) is satisfied with $s = 4$ and $q = 6$. □

Remark 7.38 (On the Proof of Theorem 7.35)

- Using the decomposition (6.64) instead of (6.65) gives in the error bound the term $\left\| \nabla \boldsymbol{u}^h \right\|_{L^4\left(0,t;L^2(\Omega)\right)}$. However, the boundedness of this term is not known and it does not follow from the stability bounds for the finite element solution. A higher regularity, which can be assumed for the solution of the continuous problem, cannot simply be assumed for the finite element solution.
- The Stokes projection $I_{\mathrm{St}}^h \boldsymbol{u}$ is defined in (4.54) with $p = 0$. If $\boldsymbol{u}$ is sufficiently smooth, the right-hand side of the strong form of (4.54) is just $-\Delta \boldsymbol{u}$. Multiplying the strong form of (4.54) with ν and inserting the strong form of the Navier–Stokes equations (7.1) gives an equivalent problem for the Stokes projection: Find $\left(I_{\mathrm{St}}^h \boldsymbol{u}, I_{\mathrm{St}}^h p\right) \in V^h \times Q^h$ such that for all $\left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h$

$$\begin{aligned} \left(\nu \nabla I_{\mathrm{St}}^h \boldsymbol{u}, \nabla \boldsymbol{v}^h\right) + \left(\nabla \cdot \boldsymbol{v}^h, I_{\mathrm{St}}^h p\right) &= \left(\boldsymbol{f} - \partial_t \boldsymbol{u} - (\boldsymbol{u} \cdot \nabla)\, \boldsymbol{u} - \nabla p, \boldsymbol{v}^h\right), \\ \left(\nabla \cdot I_{\mathrm{St}}^h \boldsymbol{u}, q^h\right) &= 0. \end{aligned} \tag{7.57}$$

This form can be found in de Frutos et al. (2016b).

Since the viscous term in (7.57) contains ν, the error estimates (4.45), (4.46), and (4.47) of the scaled Stokes equations apply for (7.57). Because $(\boldsymbol{u}, 0)$ is the solution of the continuous counterpart of (7.57), the error bounds for the velocity do not depend directly on inverse powers of the viscosity.

- Using a different projection or interpolant than $I_{\mathrm{St}}^h \boldsymbol{u}$, the term $(\nu\nabla\boldsymbol{\eta}, \nabla\boldsymbol{v}^h)$ will generally not vanish in (7.48). Then, this term is estimated with the Cauchy–Schwarz inequality (A.10) and Young's inequality (A.5).

□

Remark 7.39 (Decreasing the Dependency on ν by Using Higher Regularity Assumptions for the Velocity) Estimate (7.51) leads to ν^{-3} in the exponential of the error bound (7.44). Assuming in particular higher regularity in space, e.g., $\boldsymbol{u} \in W^{1,\infty}(\Omega)$ for the considered time t, like in Burman and Fernández (2007) and Arndt et al. (2015), this term can be estimated as follows, using (6.39) and Young's inequality (A.5),

$$
\begin{aligned}
&\left|n_{\text{skew}}\left(\boldsymbol{\phi}^h, \boldsymbol{u}, \boldsymbol{\phi}^h\right)\right| \qquad (7.58)\\
&\leq \frac{1}{2}\left(\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 \|\nabla \boldsymbol{u}\|_{L^\infty(\Omega)} + \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \|\boldsymbol{u}\|_{L^\infty(\Omega)}\right)\\
&\leq \left(\frac{1}{2}\|\nabla \boldsymbol{u}\|_{L^\infty(\Omega)} + \frac{4}{\nu}\|\boldsymbol{u}\|_{L^\infty(\Omega)}^2\right)\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 + \frac{\nu}{16}\left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 .
\end{aligned}
$$

The last term can be absorbed in the left-hand side of (7.49). With the assumptions $\nabla\boldsymbol{u} \in L^1(0,T;L^\infty(\Omega))$ and $\boldsymbol{u} \in L^2(0,T;L^\infty(\Omega))$, Gronwall's lemma can be applied and the first term enters the exponential that now depends on ν^{-1}.

If higher regularity of the solution is assumed, it was observed for the Smagorinsky turbulence model already in John and Layton (2002) that the dependency of finite element error bounds on ν can be reduced, compare Theorem 8.120. □

Remark 7.40 (Exponential Without Explicit Viscosity in the Case of Higher Regularity and $V_{\mathrm{div}}^h \subset V_{\mathrm{div}}$) Using a weakly divergence-free and inf-sup stable pair of finite element spaces, like the Scott–Vogelius space on barycentric-refined grids, see Remarks 3.135 and 3.136, the finite element error analysis can be performed for $n_{\mathrm{conv}}(\cdot,\cdot,\cdot)$ instead of $n_{\mathrm{skew}}(\cdot,\cdot,\cdot)$, since $n_{\mathrm{conv}}(\boldsymbol{u}^h, \boldsymbol{v}^h, \boldsymbol{v}^h) = 0$ for all $\boldsymbol{v}^h \in V^h$ in this case, compare (6.28). If $\boldsymbol{u} \in W^{1,\infty}(\Omega)$, one obtains with (6.38)

$$
\left|n_{\text{conv}}\left(\boldsymbol{\phi}^h, \boldsymbol{u}, \boldsymbol{\phi}^h\right)\right| \leq \|\nabla\boldsymbol{u}\|_{L^\infty(\Omega)} \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 .
$$

Under the assumption that $\nabla\boldsymbol{u} \in L^1(0,T;L^\infty(\Omega))$, Gronwall's lemma can be applied, leading to an argument of the exponential that does not depend explicitly on inverse powers of ν. □

Example 7.41 (Simulations with Analytical Solution Supporting Error Estimate (7.44)) For supporting error estimate (7.44), simulations have to be performed where the Navier–Stokes equations need to be discretized in space and time. Thus, not only a discretization error with respect to space is committed, but also with respect to time, e.g., compare Theorem 7.72. To support estimates that were derived for the time-continuous case, one has to perform simulations where the discretization error in space dominates and the error with respect to time is negligible. To this end, one can apply a time stepping scheme, if possible of higher order, with a very small time step.

Simulations were performed for Example D.10 with $\Omega = (-1,1)\times(-1,1)\times(0,2)$, $\alpha = \pi/4$, and $\gamma = \pi/2$ in the time interval $[0, 0.1]$. As discretization in space, the Taylor–Hood pair Q_2/Q_1 was used and the coarsest grid (level 0) consisted of eight cubes. The skew-symmetric form $n_{\text{skew}}(\cdot,\cdot,\cdot)$ of the convective term was utilized. For discretizing the time derivative, the Crank–Nicolson scheme, see Example 7.49, was applied with an equidistant time step $\Delta t = 0.001$.

The behavior of different errors with respect to the spatial refinement is shown in Fig. 7.1. It can be seen that the $L^2(\Omega)$ error of the velocity at the final time converges of third order for large values of ν. Thus, estimate (7.44) seems to be suboptimal with respect to this error. Smaller values of ν lead to larger errors. A second order convergence can be observed for the error of the velocity gradient in

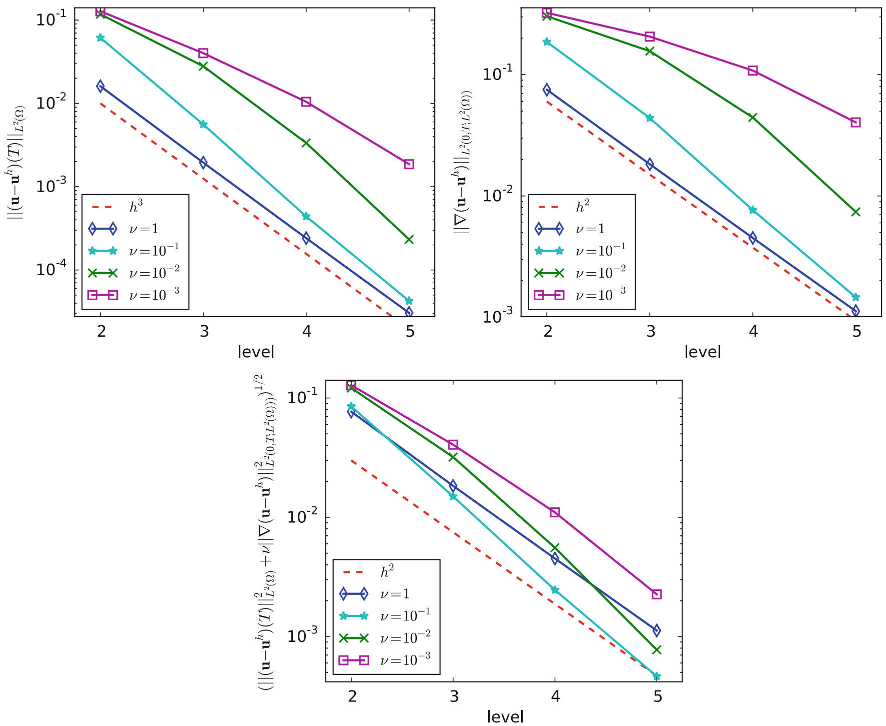

Fig. 7.1 Example 7.41. Convergence of both parts of the error from the left-hand side of estimate (7.44) and of the sum of these errors

$L^2\left(0,T;L^2(\Omega)\right)$. Again, smaller values of ν cause larger errors. Estimate (7.44) holds for a linear combination of the hitherto discussed errors, where the error for the gradient is scaled with $\nu^{1/2}$. It can be seen that this linear combination converges of second order, as predicted by the error bound, and that there are only small differences for different values of ν. Thus, the error bound is, at least for this example, (much) too pessimistic concerning the dependency of the studied error on the viscosity. □

Theorem 7.42 (Finite Element Error Estimate for the Pressure of the Continuous-in-Time Galerkin Finite Element Method) *Let the assumptions of Theorem 7.35 be satisfied and assume in addition*

$$\nabla \boldsymbol{u} \in L^\infty\left(0,T;L^2(\Omega)\right) \tag{7.59}$$

and that the family of triangulations is quasi-uniform. Then, the following error estimate holds for all $t \in (0,T]$

$$\begin{aligned}
&\left\|p-p^h\right\|^2_{L^2\left(0,t;L^2(\Omega)\right)} \\
&\le C\left(1+\frac{1}{\beta^h_{\mathrm{is}}}\right)^2 \inf_{q^h\in L^2(0,t;Q^h)} \left\|p-q^h\right\|^2_{L^2\left(0,t;L^2(\Omega)\right)} \\
&\quad + C\left(\frac{1}{\beta^h_{\mathrm{is}}}\right)^2 \Bigg[\nu^2 \left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|^2_{L^2\left(0,t;L^2(\Omega)\right)} \\
&\quad + \left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{L^\infty\left(0,t;L^2(\Omega)\right)} \left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2\left(0,t;L^2(\Omega)\right)} \left\|\nabla \boldsymbol{u}\right\|^2_{L^4\left(0,t;L^2(\Omega)\right)} \\
&\quad + \left\|\boldsymbol{u}^h\right\|_{L^\infty\left(0,t;L^2(\Omega)\right)} \left\|\nabla \boldsymbol{u}^h\right\|_{L^\infty\left(0,t;L^2(\Omega)\right)} \left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|^2_{L^2\left(0,t;L^2(\Omega)\right)} \\
&\quad + \left\|\partial_t\left(\boldsymbol{u}-I^h_{\mathrm{St}}\boldsymbol{u}\right)\right\|^2_{L^2(0,t;V')}\Bigg].
\end{aligned} \tag{7.60}$$

Proof Consider $t \in (0,T]$. For this time, the proof starts in the same way as for steady-state problems by decomposing the pressure error and applying the discrete inf-sup condition (3.51), compare the proof of Theorem 6.30. One arrives at a formula of type (6.68) with the numerator

$$\begin{aligned}
&\left(\partial_t\left(\boldsymbol{u}-\boldsymbol{u}^h\right),\boldsymbol{v}^h\right) + \nu\left(\nabla(\boldsymbol{u}-\boldsymbol{u}^h),\nabla \boldsymbol{v}^h\right) \\
&\quad + n_{\mathrm{skew}}\left(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}^h\right) - n_{\mathrm{skew}}\left(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{v}^h\right) - \left(\nabla\cdot \boldsymbol{v}^h, p-q^h\right).
\end{aligned} \tag{7.61}$$

The viscous term and the term with the pressure are bounded in the usual way by using the estimate for the dual pairing or the Cauchy–Schwarz inequality (A.10) and (3.41)

$$\nu\left(\nabla(\boldsymbol{u}-\boldsymbol{u}^h),\nabla\boldsymbol{v}^h\right) \le \nu\left\|\nabla(\boldsymbol{u}-\boldsymbol{u}^h)\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)}, \tag{7.62}$$

$$\left(\nabla\cdot\boldsymbol{v}^h, p-q^h\right) \le \left\|p-q^h\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)}. \tag{7.63}$$

The second factors cancel with the denominator of the discrete inf-sup condition. Next, the estimate of the convective terms will be discussed, which is based on the decomposition

$$\begin{aligned}
&n_{\text{skew}}\left(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}^h\right) - n_{\text{skew}}\left(\boldsymbol{u}^h,\boldsymbol{u}^h,\boldsymbol{v}^h\right)\\
&\quad = n_{\text{skew}}\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{u},\boldsymbol{v}^h\right) + n_{\text{skew}}\left(\boldsymbol{u}^h,\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{v}^h\right).
\end{aligned} \tag{7.64}$$

Applying estimate (6.41) for $s = 1/2$ gives for the first term

$$\begin{aligned}
&\left|n_{\text{skew}}\left(\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{u},\boldsymbol{v}^h\right)\right|\\
&\le C\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{L^2(\Omega)}^{1/2}\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}^{1/2}\left\|\nabla\boldsymbol{u}\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)}.
\end{aligned} \tag{7.65}$$

Again, the last term cancels with the denominator of the discrete inf-sup condition. For considering the pressure error in $L^2\left(0,t;L^2(\Omega)\right)$, the other terms have to be squared and integrated in time. Using the Cauchy–Schwarz inequality yields

$$\begin{aligned}
&\int_0^t \left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{L^2(\Omega)}\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{u}\right\|_{L^2(\Omega)}^2\,d\tau\\
&\le \left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{L^\infty\left(0,t;L^2(\Omega)\right)}\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2\left(0,t;L^2(\Omega)\right)}\left\|\nabla\boldsymbol{u}\right\|_{L^4\left(0,t;L^2(\Omega)\right)}^2.
\end{aligned}$$

The first two factors of this bound are estimated in (7.44) and the third factor is assumed to be bounded in (7.43). Applying the same approach to the second term in (7.64) yields

$$\begin{aligned}
&\left|n_{\text{skew}}\left(\boldsymbol{u}^h,\boldsymbol{u}-\boldsymbol{u}^h,\boldsymbol{v}^h\right)\right|\\
&\le C\left\|\boldsymbol{u}^h\right\|_{L^2(\Omega)}^{1/2}\left\|\nabla\boldsymbol{u}^h\right\|_{L^2(\Omega)}^{1/2}\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)}.
\end{aligned} \tag{7.66}$$

Canceling $\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)}$, taking the square, and integrating in $(0,t)$ leads to

$$\begin{aligned}
&\int_0^t \left\|\boldsymbol{u}^h\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{u}^h\right\|_{L^2(\Omega)}\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}^2\,d\tau\\
&\le \left\|\boldsymbol{u}^h\right\|_{L^\infty\left(0,t;L^2(\Omega)\right)}\int_0^t \left\|\nabla\boldsymbol{u}^h\right\|_{L^2(\Omega)}\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}^2\,d\tau.
\end{aligned}$$

The first term is bounded by the stability estimate (7.42). Applying the triangle inequality with the Stokes projection $I_{\mathrm{St}}^h \boldsymbol{u}$ at time t with $p = 0$, (4.57) with $\left(\boldsymbol{v}^h, q^h\right) = (\boldsymbol{0}, 0)$, the inverse estimate (C.35), and the notation of the proof of Theorem 7.35 gives

$$\begin{aligned} \left\|\nabla \boldsymbol{u}^h\right\|_{L^2(\Omega)} &\le \left\|\nabla\left(\boldsymbol{u} - I_{\mathrm{St}}^h \boldsymbol{u}\right)\right\|_{L^2(\Omega)} + \left\|\nabla\left(I_{\mathrm{St}}^h \boldsymbol{u} - \boldsymbol{u}^h\right)\right\|_{L^2(\Omega)} + \|\nabla \boldsymbol{u}\|_{L^2(\Omega)} \\ &\le C\|\nabla \boldsymbol{u}\|_{L^2(\Omega)} + Ch^{-1}\left\|I_{\mathrm{St}}^h \boldsymbol{u} - \boldsymbol{u}^h\right\|_{L^2(\Omega)} + \|\nabla \boldsymbol{u}\|_{L^2(\Omega)} \\ &= C\|\nabla \boldsymbol{u}\|_{L^2(\Omega)} + Ch^{-1}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}. \end{aligned} \tag{7.67}$$

The second term is bounded in $L^\infty\left(0, t; L^2(\Omega)\right)$ by (7.56) and if the convergence is at least of first order, then the negative power of h is compensated. For the other term on the right-hand side of (7.67), one has to assume (7.59). With this assumption, also the second term on the right-hand side of (7.64) can be bounded.

Analyzing the term in (7.61) with the temporal derivative starts with the decomposition

$$\left(\partial_t\left(\boldsymbol{u} - \boldsymbol{u}^h\right), \boldsymbol{v}^h\right) = \left(\partial_t\left(\boldsymbol{u} - I_{\mathrm{St}}^h \boldsymbol{u}\right), \boldsymbol{v}^h\right) + \left(\partial_t\left(I_{\mathrm{St}}^h \boldsymbol{u} - \boldsymbol{u}^h\right), \boldsymbol{v}^h\right).$$

The first term is bounded by using the estimate of the dual pairing

$$\left(\partial_t\left(\boldsymbol{u} - I_{\mathrm{St}}^h \boldsymbol{u}\right), \boldsymbol{v}^h\right) \le \left\|\partial_t\left(\boldsymbol{u} - I_{\mathrm{St}}^h \boldsymbol{u}\right)\right\|_{V'}\left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)}. \tag{7.68}$$

Bounding the other term starts in the same way. Using the notation of the proof of Theorem 7.35, one gets

$$\left(\partial_t\left(I_{\mathrm{St}}^h \boldsymbol{u} - \boldsymbol{u}^h\right), \boldsymbol{v}^h\right) \le \left\|\partial_t\left(I_{\mathrm{St}}^h \boldsymbol{u} - \boldsymbol{u}^h\right)\right\|_{V'}\left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)} = \left\|\partial_t \boldsymbol{\phi}^h\right\|_{V'}\left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)}. \tag{7.69}$$

For performing the estimate of the first factor, the operator $A^h : V_{\mathrm{div}}^h \to V_{\mathrm{div}}^h$ defined by

$$\left(A^h \boldsymbol{v}^h, \boldsymbol{w}^h\right) = \left(\nabla \boldsymbol{v}^h, \nabla \boldsymbol{w}^h\right) \quad \forall\ \boldsymbol{v}^h, \boldsymbol{w}^h \in V_{\mathrm{div}}^h \tag{7.70}$$

is introduced. The operator A^h is a discrete Stokes operator. It is symmetric (self-adjoint) and positive definite. Taking $\boldsymbol{w}^h = \boldsymbol{v}^h$ in (7.70) gives

$$\left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)} = \left(A^h \boldsymbol{v}^h, \boldsymbol{v}^h\right)^{1/2} = \left(\left(A^h\right)^{1/2} \boldsymbol{v}^h, \left(A^h\right)^{1/2} \boldsymbol{v}^h\right)^{1/2} = \left\|\left(A^h\right)^{1/2} \boldsymbol{v}^h\right\|_{L^2(\Omega)}. \tag{7.71}$$

Setting both arguments in (7.70) to $\left(A^h\right)^{-1/2} \boldsymbol{v}^h$ leads to

$$\left\| \nabla \left(A^h\right)^{-1/2} \boldsymbol{v}^h \right\|_{L^2(\Omega)} = \left(A^h \left(A^h\right)^{-1/2} \boldsymbol{v}^h, \left(A^h\right)^{-1/2} \boldsymbol{v}^h \right)^{1/2}$$
$$= \left(\left(A^h\right)^{1/2} \boldsymbol{v}^h, \left(A^h\right)^{-1/2} \boldsymbol{v}^h \right)^{1/2} = \left\| \boldsymbol{v}^h \right\|_{L^2(\Omega)}. \quad (7.72)$$

Applying the results Ayuso et al. (2005, Lemma 3.11 and (2.15)) gives the estimate

$$\left\| \partial_t \boldsymbol{\phi}^h \right\|_{V'} \leq Ch \left\| \partial_t \boldsymbol{\phi}^h \right\|_{L^2(\Omega)} + C \left\| \left(A^h\right)^{-1/2} \partial_t \boldsymbol{\phi}^h \right\|_{L^2(\Omega)}.$$

The first term on the right-hand side can be bounded by using (7.71) and the inverse estimate (C.35)

$$Ch \left\| \partial_t \boldsymbol{\phi}^h \right\|_{L^2(\Omega)} = Ch \left\| \left(A^h\right)^{1/2} \left(A^h\right)^{-1/2} \partial_t \boldsymbol{\phi}^h \right\|_{L^2(\Omega)}$$
$$= Ch \left\| \nabla \left(\left(A^h\right)^{-1/2} \partial_t \boldsymbol{\phi}^h \right) \right\|_{L^2(\Omega)} \leq C \left\| \left(A^h\right)^{-1/2} \partial_t \boldsymbol{\phi}^h \right\|_{L^2(\Omega)},$$

such that

$$\left\| \partial_t \boldsymbol{\phi}^h \right\|_{V'} \leq C \left\| \left(A^h\right)^{-1/2} \partial_t \boldsymbol{\phi}^h \right\|_{L^2(\Omega)}. \quad (7.73)$$

Bounding the right-hand side of (7.73) uses the error equation (7.48) with the test function $\boldsymbol{v}^h = \left(A^h\right)^{-1} \partial_t \boldsymbol{\phi}^h \in V_{\text{div}}^h$. The individual terms will be considered. One obtains

$$\left(\partial_t \boldsymbol{\phi}^h, \left(A^h\right)^{-1} \partial_t \boldsymbol{\phi}^h \right) = \left(\left(A^h\right)^{-1/2} \partial_t \boldsymbol{\phi}^h, \left(A^h\right)^{-1/2} \partial_t \boldsymbol{\phi}^h \right) = \left\| \left(A^h\right)^{-1/2} \partial_t \boldsymbol{\phi}^h \right\|_{L^2(\Omega)}^2. \quad (7.74)$$

Using the estimate of the dual pairing and (7.72), one gets

$$\left(\partial_t \boldsymbol{\eta}, \left(A^h\right)^{-1} \partial_t \boldsymbol{\phi}^h \right) \leq \left\| \partial_t \boldsymbol{\eta} \right\|_{V'} \left\| \nabla \left(\left(A^h\right)^{-1} \partial_t \boldsymbol{\phi}^h \right) \right\|_{L^2(\Omega)}$$
$$= \left\| \partial_t \boldsymbol{\eta} \right\|_{V'} \left\| \nabla \left(\left(A^h\right)^{-1/2} \left(A^h\right)^{-1/2} \partial_t \boldsymbol{\phi}^h \right) \right\|_{L^2(\Omega)}$$
$$= \left\| \partial_t \left(\boldsymbol{u} - I_{\text{St}}^h \boldsymbol{u} \right) \right\|_{V'} \left\| \left(A^h\right)^{-1/2} \partial_t \boldsymbol{\phi}^h \right\|_{L^2(\Omega)}. \quad (7.75)$$

The viscous term is bounded by using the definition (7.70) of the discrete Stokes operator, the Cauchy–Schwarz inequality, and (7.71)

$$\begin{aligned}
&\left(\nu\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right), \nabla\left(A^h\right)^{-1}\partial_t\boldsymbol{\phi}^h\right)\\
&= \nu\left(A^h\left(\boldsymbol{u}-\boldsymbol{u}^h\right), \left(A^h\right)^{-1}\partial_t\boldsymbol{\phi}^h\right) = \nu\left(\left(A^h\right)^{1/2}\left(\boldsymbol{u}-\boldsymbol{u}^h\right), \left(A^h\right)^{-1/2}\partial_t\boldsymbol{\phi}^h\right)\\
&\le \nu\left\|\left(A^h\right)^{1/2}\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}\left\|\left(A^h\right)^{-1/2}\partial_t\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}\\
&= \nu\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}\left\|\left(A^h\right)^{-1/2}\partial_t\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}. \qquad (7.76)
\end{aligned}$$

The estimate of the pressure term applies the Cauchy–Schwarz inequality, (3.41), and (7.72), where the term with the test function is handled as in estimate (7.75),

$$\begin{aligned}
\left(\nabla\cdot\left(\left(A^h\right)^{-1}\partial_t\boldsymbol{\phi}^h\right), p-q^h\right) &\le \left\|p-q^h\right\|_{L^2(\Omega)}\left\|\nabla\left(\left(A^h\right)^{-1}\partial_t\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)}\\
&= \left\|p-q^h\right\|_{L^2(\Omega)}\left\|\left(A^h\right)^{-1/2}\partial_t\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}. \qquad (7.77)
\end{aligned}$$

Applying for the nonlinear term the same techniques that led to estimates (7.65) and (7.66), and treating the term with the test function in the same way as in (7.75) gives

$$\begin{aligned}
&n_{\text{skew}}\left(\boldsymbol{u},\boldsymbol{u},\left(A^h\right)^{-1}\partial_t\boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\boldsymbol{u}^h,\boldsymbol{u}^h,\left(A^h\right)^{-1}\partial_t\boldsymbol{\phi}^h\right)\\
&\le C\Big(\left\|\boldsymbol{u}-\boldsymbol{u}^h\right\|_{L^2(\Omega)}^{1/2}\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}^{1/2}\left\|\nabla\boldsymbol{u}\right\|_{L^2(\Omega)}\\
&\quad+\left\|\boldsymbol{u}^h\right\|_{L^2(\Omega)}^{1/2}\left\|\nabla\boldsymbol{u}^h\right\|_{L^2(\Omega)}^{1/2}\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)\right\|_{L^2(\Omega)}\Big)\left\|\left(A^h\right)^{-1/2}\partial_t\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}. \qquad (7.78)
\end{aligned}$$

The second factors of estimates (7.75)–(7.78) cancel with one of the factors of (7.74), leading to a bound for $\left\|\left(A^h\right)^{-1/2}\partial_t\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}$. This bound contains only terms that already appeared in estimates (7.62), (7.63), (7.65), (7.66), and (7.68). The bound for $\left\|\left(A^h\right)^{-1/2}\partial_t\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}$ is inserted in (7.73) and the resulting bound in (7.69), which finishes the estimate of the term with the temporal derivative. Altogether, this estimate leads to the same terms that appeared already in the bounds of the other terms of (7.61).

Collecting all estimates, i.e., estimates (7.62), (7.63), (7.65), (7.66), (7.68), and (7.69) yields for all $q^h \in Q^h$

$$
\begin{aligned}
&\left\| p^h - q^h \right\|_{L^2(\Omega)} \\
&\le \frac{C}{\beta_{\mathrm{is}}^h} \Big(\nu \left\| \nabla (\boldsymbol{u} - \boldsymbol{u}^h) \right\|_{L^2(\Omega)} + \left\| \boldsymbol{u} - \boldsymbol{u}^h \right\|_{L^2(\Omega)}^{1/2} \left\| \nabla \left(\boldsymbol{u} - \boldsymbol{u}^h\right) \right\|_{L^2(\Omega)}^{1/2} \left\| \nabla \boldsymbol{u} \right\|_{L^2(\Omega)} \\
&\quad + \left\| \boldsymbol{u}^h \right\|_{L^2(\Omega)}^{1/2} \left\| \nabla \boldsymbol{u}^h \right\|_{L^2(\Omega)}^{1/2} \left\| \nabla \left(\boldsymbol{u} - \boldsymbol{u}^h\right) \right\|_{L^2(\Omega)} + \left\| \partial_t \left(\boldsymbol{u} - I_{\mathrm{St}}^h \boldsymbol{u}\right) \right\|_{V'} \\
&\quad + \left\| p - q^h \right\|_{L^2(\Omega)} \Big). \qquad (7.79)
\end{aligned}
$$

Applying the triangle inequality

$$
\left\| p - p^h \right\|_{L^2(\Omega)} \le \left\| p - q^h \right\|_{L^2(\Omega)} + \left\| p^h - q^h \right\|_{L^2(\Omega)},
$$

inserting estimate (7.79), taking the square, and integrating on $(0, t)$ gives the error estimate (7.60) for the pressure. ■

Remark 7.43 (On Error Estimate (7.60)) The error bound on the right-hand side of (7.60) contains velocity errors that are estimated in (7.44) and terms that appear in the error bound of the velocity estimate. It follows that the order of convergence for $\left\| p - p^h \right\|_{L^2(0,t;L^2(\Omega))}$ is the same as for the velocity errors on the left-hand side of (7.44). □

Example 7.44 (Simulations with Analytical Solution Supporting Error Estimate (7.60)) Simulations with the Q_2/P_1^{disc} pair of finite element spaces were performed with Example D.7. The initial grid, level 0, was the irregular quadrilateral grid presented in Fig. 4.2. As temporal discretization, the Crank–Nicolson scheme with the equidistant time step $\Delta t = 10^{-4}$ was applied. The final time was set to be $T = 5$.

Results for the pressure error are shown in Fig. 7.2. The second order convergence can be clearly observed. In this example, the pressure errors do not depend on the viscosity. □

Remark 7.45 (Numerical Analysis for Averaged Quantities) Time-averaged or space-time-averaged quantities are of interest in many applications. So far, however, there are only very few results from the numerical analysis concerning such quantities, e.g., see John et al. (2007). □

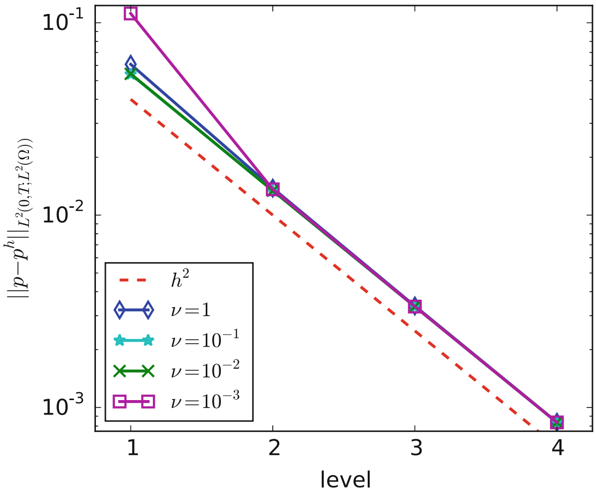

Fig. 7.2 Example 7.44. Convergence of the pressure error

7.3 Temporal Discretizations Leading to Coupled Problems

7.3.1 *θ-Schemes as Discretization in Time*

Remark 7.46 (Principal Approach of the Application of θ-Schemes) θ-schemes use the following strategy for the full discretization and linearization of (7.1):

1. *Semi-discretization of* (7.1) *in time.* The semi-discretization in time leads in each discrete time to a nonlinear system of equations of saddle point type.
2. *Variational formulation and linearization.* The nonlinear system of equations is reformulated as variational problem and the nonlinear variational problem is linearized.
3. *Discretization of the linear systems in space.* The linear system of equations arising in each step of the iteration for solving the nonlinear problem is discretized by a finite element discretization using, e.g., an inf-sup stable pair of finite element spaces.

This approach, which applies first the discretization in time and then in space, is also called method of Rothe or horizontal method of lines. The individual steps in this strategy are described in detail in this section. The other way, discretizing first in space to get an ordinary differential equation and then in time, is called (vertical) method of lines. □

Remark 7.47 (Notation) Let $t_n \in [0, T]$, then quantities at time t_n are denoted with a subscript n. The length of the time step from a discrete time t_n to the next discrete time t_{n+1} is denoted by Δt_{n+1}, i.e., $\Delta t_{n+1} = t_{n+1} - t_n$. □

Remark 7.48 (One-Step θ-Schemes for the Navier–Stokes Equations) Let $\theta \in [0, 1]$ and consider the time step from t_n to t_{n+1}. One-step θ-schemes for the Navier–Stokes equations are of the form

$$\begin{aligned} &\frac{\boldsymbol{u}_{n+1}}{\Delta t_{n+1}} + \theta\big[-\nu\Delta\boldsymbol{u}_{n+1} + (\boldsymbol{u}_{n+1}\cdot\nabla)\,\boldsymbol{u}_{n+1}\big] + \nabla p_{n+1} \\ &= \frac{\boldsymbol{u}_n}{\Delta t_{n+1}} - (1-\theta)\big[-\nu\Delta\boldsymbol{u}_n + (\boldsymbol{u}_n\cdot\nabla)\,\boldsymbol{u}_n\big] + (1-\theta)\boldsymbol{f}_n + \theta\boldsymbol{f}_{n+1}, \\ 0 &= \nabla\cdot\boldsymbol{u}_{n+1}. \end{aligned} \tag{7.80}$$

□

Example 7.49 (Explicit and Implicit Euler Scheme, Crank–Nicolson Scheme) Three well known one-step θ-schemes, the forward and backward Euler scheme and the Crank–Nicolson scheme, are obtained by an appropriate choice of θ:

- $\theta = 0$: forward or explicit Euler scheme,
- $\theta = 0.5$: Crank–Nicolson scheme,
- $\theta = 1$: backward or implicit Euler scheme.

A few remarks concerning these schemes should be given already here. More detailed comments will be provided in Remark 7.59, after the presentation of a numerical example.

- *Forward or explicit Euler scheme.* The appearance of the viscous term, which has the same form as a diffusive term, leads to a stiff ordinary differential equation with respect to time, see Hairer and Wanner (2010, p. 6). From the numerical analysis of ordinary differential equations, it is known that explicit schemes have to be used with very small time steps for stiff problems to obtain stable simulations. For the Navier–Stokes equations, the time step has to be usually so small that the simulations become very inefficient. For this reason, the forward Euler scheme and any other explicit scheme are not recommended for the discretization of the incompressible Navier–Stokes equations.
- *Backward or implicit Euler scheme.* This first order scheme is quite popular. However, it is known from the literature, e.g., from John et al. (2006a), and it will be demonstrated in Example 7.58, that the use of the backward Euler scheme in combination with higher order discretizations in space might lead to rather inaccurate results, compared with the results computed with higher order temporal discretizations.
- *Crank–Nicolson scheme, trapezoidal rule.* The Crank–Nicolson scheme is a second order scheme whose use is very popular.
- *General.* A defect in the pressure approximation was observed and analyzed in Besier and Wollner (2012) for one-step schemes on adaptive grids that change in time. Several proposals to handle this defect were studied in this paper. The final conclusion was that the best approach consists in using a scheme like the

fractional-step θ-scheme presented in Example 7.52, at least in the time step after the change of the grid.

□

Remark 7.50 (Concerning the Incompressibility Constraint)

- The temporal derivative is approximated in (7.80) at $t_n + \theta \Delta t_{n+1}$, compare (7.94). Thus, the velocity is approximated at this time and the pressure acts as a Lagrangian multiplier for this velocity. For instance, in the case of the Crank–Nicolson scheme, the computed solution approximates $\left(\boldsymbol{u}_{n+1/2}, p_{n+1/2}\right)$. If one is interested in a good approximation of the pressure at time t_{n+1}, one has to modify the pressure term in the Crank–Nicolson scheme to

$$\frac{1}{2}\left(p_{n+1} + p_n\right),$$

e.g., see Rang (2008). However, this scheme requires in the first step the pressure at the initial time, see Remark 7.29 for a discussion of this topic.
- The variational form of the time step (7.80) can be written in the product space, analogously to the form (4.3) for the Stokes equations: Find $(\boldsymbol{u}_{n+1}, p_{n+1}) \in V \times Q$ such that

$$\begin{aligned}
&\frac{1}{\Delta t_{n+1}}\left(\boldsymbol{u}_{n+1}, \boldsymbol{v}\right) + \theta\left[\nu\left(\nabla \boldsymbol{u}_{n+1}, \nabla \boldsymbol{v}\right) + \left(\left(\boldsymbol{u}_{n+1} \cdot \nabla\right) \boldsymbol{u}_{n+1}, \boldsymbol{v}\right)\right] \\
&\quad + \left(\nabla \cdot \boldsymbol{u}_{n+1}, q\right) - \left(\nabla \cdot \boldsymbol{v}, p_{n+1}\right) \\
&= \frac{1}{\Delta t_{n+1}}\left(\boldsymbol{u}_n, \boldsymbol{v}\right) - (1-\theta)\left[\nu\left(\nabla \boldsymbol{u}_n, \nabla \boldsymbol{v}\right) + \left(\left(\boldsymbol{u}_n \cdot \nabla\right) \boldsymbol{u}_n, \boldsymbol{v}\right)\right] \\
&\quad + (1-\theta)\langle \boldsymbol{f}_n, \boldsymbol{v}\rangle_{V',V} + \theta\langle \boldsymbol{f}_{n+1}\rangle_{V',V},
\end{aligned}$$

for $(\boldsymbol{v}, q) \in V \times Q$ or equivalently

$$\begin{aligned}
&\left(\boldsymbol{u}_{n+1}, \boldsymbol{v}\right) + \theta \Delta t_{n+1}\left[\nu\left(\nabla \boldsymbol{u}_{n+1}, \nabla \boldsymbol{v}\right) + \left(\left(\boldsymbol{u}_{n+1} \cdot \nabla\right) \boldsymbol{u}_{n+1}, \boldsymbol{v}\right)\right] \\
&\quad + \Delta t_{n+1}\left(\nabla \cdot \boldsymbol{u}_{n+1}, q\right) - \Delta t_{n+1}\left(\nabla \cdot \boldsymbol{v}, p_{n+1}\right) \\
&= \left(\boldsymbol{u}_n, \boldsymbol{v}\right) - (1-\theta)\Delta t_{n+1}\left[\nu\left(\nabla \boldsymbol{u}_n, \nabla \boldsymbol{v}\right) + \left(\left(\boldsymbol{u}_n \cdot \nabla\right) \boldsymbol{u}_n, \boldsymbol{v}\right)\right] \\
&\quad + (1-\theta)\Delta t_{n+1}\langle \boldsymbol{f}_n, \boldsymbol{v}\rangle_{V',V} + \theta \Delta t_{n+1}\langle \boldsymbol{f}_{n+1}, \boldsymbol{v}\rangle_{V',V}.
\end{aligned} \tag{7.81}$$

Using in (7.81) $(\boldsymbol{v}, q) = (\boldsymbol{u}, p)$, one finds that the terms with the divergence cancel, as it was used in the analysis for steady-state problems, e.g., in the proof of Lemma 6.21. Thus, (7.81) possesses in this respect the same properties as the formulation for steady-state problems. The corresponding strong formulation of (7.81) reads as follows

$$\begin{aligned}
&\boldsymbol{u}_{n+1} + \theta \Delta t_{n+1}\big[-\nu \Delta \boldsymbol{u}_{n+1} + (\boldsymbol{u}_{n+1}\cdot\nabla)\,\boldsymbol{u}_{n+1}\big] + \Delta t_{n+1}\nabla p_{n+1} \\
&\quad = \boldsymbol{u}_n - (1-\theta)\Delta t_{n+1}\big[-\nu\Delta \boldsymbol{u}_n + (\boldsymbol{u}_n\cdot\nabla)\,\boldsymbol{u}_n\big] + (1-\theta)\Delta t_{n+1}\boldsymbol{f}_n \\
&\qquad + \theta\Delta t_{n+1}\boldsymbol{f}_{n+1}, \\
&0 = \Delta t_{n+1}\nabla\cdot\boldsymbol{u}_{n+1}.
\end{aligned} \tag{7.82}$$

One can see that the divergence constraint is multiplied with Δt_{n+1}.

Of coarse, it is possible just to multiply the momentum equation of (7.80) with Δt_{n+1} and to leave the continuity equation as it is, which gives a mathematically equivalent problem to (7.81). Apart that the scaling with Δt_{n+1} was motivated in this remark, the numerical solution of problem (7.81) with the scaled divergence is more efficient in our experience.

In all simulations of time-dependent problems presented in this monograph, the divergence constraint was scaled with the time step as in (7.81) or (7.82).

□

Remark 7.51 (General θ-Scheme for the Navier–Stokes Equations) The time step of a general θ-scheme for the Navier–Stokes equations (7.1) has the form

$$\begin{aligned}
&\frac{\boldsymbol{u}_{k+1}}{\Delta t_{n+1}} + \theta_1\big[-\nu\Delta\boldsymbol{u}_{k+1} + (\boldsymbol{u}_{k+1}\cdot\nabla)\,\boldsymbol{u}_{k+1}\big] + \frac{\Delta t_{k+1}}{\Delta t_{n+1}}\nabla p_{k+1} \\
&\qquad = \frac{\boldsymbol{u}_k}{\Delta t_{n+1}} - \theta_2\big[-\nu\Delta\boldsymbol{u}_k + (\boldsymbol{u}_k\cdot\nabla)\,\boldsymbol{u}_k\big] + \theta_3\boldsymbol{f}_k + \theta_4\boldsymbol{f}_{k+1}, \\
&\Delta t_{k+1}\nabla\cdot\boldsymbol{u}_{k+1} = 0,
\end{aligned} \tag{7.83}$$

with the parameters $\theta_1,\ldots,\theta_4$. Note the different indices k and n. □

Example 7.52 (Fractional-Step θ-Scheme) The fractional-step θ-scheme, developed in Bristeau et al. (1987), is obtained by three steps of form (7.83). There exist two variants of this scheme. The two variants, FS0 and FS1, are presented in Table 7.1, where

$$\theta = 1 - \frac{\sqrt{2}}{2}, \quad \tilde\theta = 1 - 2\theta, \quad \tau = \frac{\tilde\theta}{1-\theta}, \quad \eta = 1-\tau.$$

A fractional-step θ-scheme is a clever combination of three first order one-step schemes to achieve a strongly A-stable second order scheme.

FS1 requires the evaluation of $\boldsymbol{f}$ only at the times t_n and $t_{n+1} - \theta\Delta t_{n+1}$ whereas FS0 needs the evaluation of $\boldsymbol{f}$ in addition at $t_n + \theta\Delta t_{n+1}$ and at t_{n+1}. Both variants are second order schemes but FS1 does not integrate second order polynomials (with respect to t) exactly. However, most other fundamental properties, like stability, are the same for both variants. Results for the two-dimensional Navier–Stokes equations, John et al. (2006a), show that FS0 is often considerable more accurate

Table 7.1 The two variants of the fractional-step θ-schemes

	θ_1	θ_2	θ_3	θ_4	t_k	t_{k+1}	Δt_{k+1}	Order
FS0	$\tau\theta$	$\eta\theta$	$\eta\theta$	$\tau\theta$	t_n	$t_n + \theta\Delta t_{n+1}$	$\theta\Delta t_{n+1}$	2
	$\eta\tilde{\theta}$	$\tau\tilde{\theta}$	$\tau\tilde{\theta}$	$\eta\tilde{\theta}$	$t_n + \theta\Delta t_{n+1}$	$t_{n+1} - \theta\Delta t_{n+1}$	$\tilde{\theta}\Delta t_{n+1}$	
	$\tau\theta$	$\eta\theta$	$\eta\theta$	$\tau\theta$	$t_{n+1} - \theta\Delta t_{n+1}$	t_{n+1}	$\theta\Delta t_{n+1}$	
FS1	$\tau\theta$	$\eta\theta$	θ	0	t_n	$t_n + \theta\Delta t_{n+1}$	$\theta\Delta t_{n+1}$	2
	$\eta\tilde{\theta}$	$\tau\tilde{\theta}$	0	$\tilde{\theta}$	$t_n + \theta\Delta t_{n+1}$	$t_{n+1} - \theta\Delta t_{n+1}$	$\tilde{\theta}\Delta t_{n+1}$	
	$\tau\theta$	$\eta\theta$	θ	0	$t_{n+1} - \theta\Delta t_{n+1}$	t_{n+1}	$\theta\Delta t_{n+1}$	

Table 7.2 Parameters for one-step θ-schemes written as general θ-scheme (7.83)

	θ_1	θ_2	θ_3	θ_4	t_k	t_{k+1}	Δt_{k+1}	Order
Forward Euler scheme	0	1	1	0	t_n	t_{n+1}	Δt_{n+1}	1
Backward Euler scheme (BWE)	1	0	0	1	t_n	t_{n+1}	Δt_{n+1}	1
Crank–Nicolson scheme (CN)	0.5	0.5	0.5	0.5	t_n	t_{n+1}	Δt_{n+1}	2

than FS1. Thus, if the evaluation of the right-hand side is not very expensive, FS0 should be preferred. □

Remark 7.53 (Representation of θ-Schemes by One Formula) Note that the one-step θ-scheme (7.80) can be written as general θ-scheme (7.83) with appropriate parameters, see Table 7.2. Hence, (7.83) enables the implementation of all presented schemes by a single formula and the choice between the schemes by setting four parameters.

All further issues for θ-schemes presented in this section will be discussed for the general θ-scheme (7.83). □

Remark 7.54 (Variational Formulation) The solution of (7.83) will be approximated by a finite element method, with the basis of the finite element method being a variational formulation of (7.83). The derivation of the variational problem is performed in the usual way by multiplying the equations in (7.83) with test functions, integrating on Ω, and applying integration by parts, which gives after scaling the whole problem with Δt_{n+1} the following problem: Find $(\boldsymbol{u}_{k+1}, p_{k+1}) \in V \times Q$ such that for all $(\boldsymbol{v}, q) \in V \times Q$

$$\begin{aligned}
(\boldsymbol{u}_{k+1}, \boldsymbol{v}) + \theta_1 \Delta t_{n+1} &\left[(\nu \nabla \boldsymbol{u}_{k+1}, \nabla \boldsymbol{v}) + ((\boldsymbol{u}_{k+1} \cdot \nabla)\, \boldsymbol{u}_{k+1}, \boldsymbol{v})\right] \\
&- \Delta t_{k+1}\, (\nabla \cdot \boldsymbol{v}, p_{k+1}) \\
&= (\boldsymbol{u}_k, \boldsymbol{v}) - \theta_2 \Delta t_{n+1} [(2\nu \nabla \boldsymbol{u}_k, \nabla \boldsymbol{v}) + ((\boldsymbol{u}_k \cdot \nabla)\, \boldsymbol{u}_k, \boldsymbol{v})] \\
&\quad + \theta_3 \Delta t_{n+1} \langle \boldsymbol{f}_k, \boldsymbol{v} \rangle_{V',V} + \theta_4 \Delta t_{n+1} \langle \boldsymbol{f}_{k+1}, \boldsymbol{v} \rangle_{V',V}, \\
0 &= \Delta t_{k+1}\, (\nabla \cdot \boldsymbol{u}_{k+1}, q)\,.
\end{aligned} \tag{7.84}$$

Of course, any other form of the convective term discussed in Sect. 6.1.2 can be used. For brevity, the approach will be described only for the convective form of the convective term. □

Remark 7.55 (Linearization of the Variational Form) The nonlinear system (7.84) is solved iteratively starting with an initial guess $\left(\boldsymbol{u}_{k+1}^0, p_{k+1}^0\right)$. The nonlinear convective term has to be linearized, where the same linearizations as for the steady-state Navier–Stokes equations can be applied, see Sect. 6.3.

Given a known velocity field $\boldsymbol{u}_{k+1}^{(m)}$, the fixed point iteration or Picard iteration uses the approximation

$$\left(\boldsymbol{u}_{k+1}^{(m+1)} \cdot \nabla\right) \boldsymbol{u}_{k+1}^{(m+1)} \approx \left(\boldsymbol{u}_{k+1}^{(m)} \cdot \nabla\right) \boldsymbol{u}_{k+1}^{(m+1)}.$$

Then, the fixed point iteration for solving (7.84) has the following form: Given $\left(\boldsymbol{u}_{k+1}^{(m)}, p_{k+1}^{(m)}\right)$, the iterate $\left(\boldsymbol{u}_{k+1}^{(m+1)}, p_{k+1}^{(m+1)}\right)$ is computed by solving

$$\begin{aligned}
\left(\boldsymbol{u}_{k+1}^{(m+1)}, \boldsymbol{v}\right) &+ \theta_1 \Delta t_{n+1} \left[\left(\nu \nabla \boldsymbol{u}_{k+1}^{(m+1)}, \nabla \boldsymbol{v}\right) + \left(\left(\boldsymbol{u}_{k+1}^{(m)} \cdot \nabla\right) \boldsymbol{u}_{k+1}^{(m+1)}, \boldsymbol{v}\right)\right] \\
&- \Delta t_{k+1} \left(\nabla \cdot \boldsymbol{v}, p_{k+1}^{(m+1)}\right) \\
&= (\boldsymbol{u}_k, \boldsymbol{v}) + \theta_3 \Delta t_{n+1} \langle \boldsymbol{f}_k, \boldsymbol{v}\rangle_{V',V} + \theta_4 \Delta t_{n+1} \langle \boldsymbol{f}_{k+1}, \boldsymbol{v}\rangle_{V',V} \qquad (7.85) \\
&- \theta_2 \Delta t_{n+1} \left[\, (\nu \nabla \boldsymbol{u}_k, \nabla \boldsymbol{v}) + ((\boldsymbol{u}_k \cdot \nabla)\, \boldsymbol{u}_k, \boldsymbol{v})\, \right], \\
0 &= \Delta t_{k+1} \left(\nabla \cdot \boldsymbol{u}_{k+1}^{(m+1)}, q\right) \quad \forall\ (\boldsymbol{v}, q) \in V \times Q,
\end{aligned}$$

$m = 0, 1, 2, \ldots$. Equation (7.85) are of Oseen type, see (5.2). The right-hand side of this equation does not change during the iteration.

For Newton's method, the nonlinear convective term is linearized as follows

$$\left(\boldsymbol{u}_{k+1}^{(m+1)} \cdot \nabla\right) \boldsymbol{u}_{k+1}^{(m+1)} \approx \left(\boldsymbol{u}_{k+1}^{(m)} \cdot \nabla\right) \boldsymbol{u}_{k+1}^{(m+1)} + \left(\boldsymbol{u}_{k+1}^{(m+1)} \cdot \nabla\right) \boldsymbol{u}_{k+1}^{(m)} - \left(\boldsymbol{u}_{k+1}^{(m)} \cdot \nabla\right) \boldsymbol{u}_{k+1}^{(m)},$$

such that the iteration is of the form: Given $\left(\boldsymbol{u}_{k+1}^{(m)}, p_{k+1}^{(m)}\right)$, the iterate $\left(\boldsymbol{u}_{k+1}^{(m+1)}, p_{k+1}^{(m+1)}\right)$ is computed by solving

$$\begin{aligned}
\left(\boldsymbol{u}_{k+1}^{(m+1)}, \boldsymbol{v}\right) &+ \theta_1 \Delta t_{n+1} \Big[\left(\nu \nabla \boldsymbol{u}_{k+1}^{(m+1)}, \nabla \boldsymbol{v}\right) + \left(\left(\boldsymbol{u}_{k+1}^{(m)} \cdot \nabla\right) \boldsymbol{u}_{k+1}^{(m+1)}, \boldsymbol{v}\right) \\
&+ \left(\left(\boldsymbol{u}_{k+1}^{(m+1)} \cdot \nabla\right) \boldsymbol{u}_{k+1}^{(m)}, \boldsymbol{v}\right) \Big] - \Delta t_{k+1} \left(\nabla \cdot \boldsymbol{v}, p_{k+1}^{(m+1)}\right) \\
&= (\boldsymbol{u}_k, \boldsymbol{v}) + \theta_3 \Delta t_{n+1} \langle \boldsymbol{f}_k, \boldsymbol{v}\rangle_{V',V} + \theta_4 \Delta t_{n+1} \langle \boldsymbol{f}_{k+1}, \boldsymbol{v}\rangle_{V',V} \qquad (7.86) \\
&+ \theta_1 \Delta t_{n+1} \left(\left(\boldsymbol{u}_{k+1}^{(m)} \cdot \nabla\right) \boldsymbol{u}_{k+1}^{(m)}, \boldsymbol{v}\right)
\end{aligned}$$

$$-\theta_2 \Delta t_{n+1}\big[\,(\nu\nabla \boldsymbol{u}_k, \nabla \boldsymbol{v}) + ((\boldsymbol{u}_k\cdot\nabla)\,\boldsymbol{u}_k, \boldsymbol{v})\,\big],$$

$$0 = \Delta t_{k+1}\left(\nabla\cdot \boldsymbol{u}_{k+1}^{(m+1)}, q\right) \quad \forall\ (\boldsymbol{v}, q) \in V\times Q,$$

$m = 0, 1, 2, \ldots$.

The initial guess can be chosen to be the solution at the previous time step $\left(\boldsymbol{u}_{k+1}^0, p_{k+1}^0\right) = (\boldsymbol{u}_k, p_k)$ or some extrapolation from more than one previous times.

The numerical approximation of the Navier–Stokes equations (7.1) using the approach described in this section requires the repeated solution of linear saddle point problems of form (7.85) or (7.86) in each discrete time. □

Remark 7.56 (Comparison of the Picard Iteration and Newton's Method) For the comparison of the Picard iteration and Newton's method, the same statements as given for the stationary Navier–Stokes equations in Sect. 6.3, e.g., in Remarks 6.44 and 6.46, apply.

A main difference to steady-state problems is the availability of a good initial iterate in time-dependent problems by using the solution of the previous discrete time or even an extrapolation from several previous solutions. Thus, the number of iterations for computing the solution in one time step is usually considerably smaller than the number of iterations needed for steady-state problems. The experience is that the shorter the time step, the smaller the number of iterations becomes.

Using an approach with an iterative solver and an inexact solution of the linear systems of equations, the numerical studies at a three-dimensional flow around a cylinder in John (2006) revealed that for this example, the Picard iteration was considerably more efficient then Newton's method. It will be discussed in Remark 9.5 that for problems in three dimensions usually iterative solvers for the linear systems of equations should be applied, compare also Example 7.57. □

Example 7.57 (Picard Iteration vs. Newton's Method) This example considers a three-dimensional problem with analytic solution, see Example D.10. The viscosity $\nu = 10^{-3}$ was chosen for the simulations. A hexahedral grid consisting of cubes was applied for triangulating the domain (eight cubes on level 0), the Q_2/Q_1 pair of finite element spaces was used, the Crank–Nicolson scheme with the equidistant time step $\Delta t = 0.01$ in the time interval $[0, 0.1]$ was applied, and the solution of the previous discrete time was used as initial iterate.

Results for the described setup are presented in Table 7.3. The general observations are the same as for the steady-state problem, compare Example 6.47. Applying Newton's method requires generally less iterations than using the Picard method. Using a direct solver for the linear systems of equations leads to a smaller number of iterations in Newton's method.

Table 7.3 provides also some information about the time for solving one linear system of equations. The average time needed for the inexact solution in the Picard method on level 2 was taken as basis (100 % $\approx$ 3 s). The inexact solver was the FGMRES method with a coupled multigrid preconditioner, see Remark 9.6 and

Table 7.3 Example 7.57

	Level 2		Level 3	
Velocity/pressure d.o.f.	14,739/729		107,811/4913	
	Lin. systems	Rel. time/system	Lin. systems	Rel. time / system
Picard/inexact	33	100	23	824
Picard/direct	33	140	22	8989
Newton/inexact	23	101	12	804
Newton/direct	11	142	11	8932

Number of solutions of linear systems of equations for computing the solution (ten time steps) and relative computing time for solving one linear system of equations (Picard inexact as 100 % ≈ 3 s)

Table 7.4 Example 7.58

Level	Velocity	Pressure	All
3	27,232	9984	37,216
4	107,712	39,936	147,648
5	428,416	159,744	588,160

Number of degrees of freedom in space (including Dirichlet nodes)

Sect. 9.2.2. As sparse direct solver, umfpack was applied, see Remark 9.5. It can be observed that the computing times on level 2 for both solvers are of the same order. Consequently, Newton's method with direct solver was the most efficient approach. Refining the grid once leads to an increase of the number of unknowns by a factor of eight. It can be seen that the inexact solver scales almost optimally whereas the direct solver becomes very inefficient. Consequently, Newton's method with inexact solver was the fastest approach on level 3. □

Example 7.58 (Accuracy Studies at Example D.9) This example considers the two-dimensional flow around a cylinder defined in Example D.9. At the outlet, the do-nothing boundary conditions (D.30) were applied.

Results are presented for three implicit θ-schemes: the backward Euler scheme, the Crank–Nicolson scheme, and the fractional-step θ-scheme. Note that in this problem $\boldsymbol{f} = \boldsymbol{0}$ for all times such that FS0 and FS1 are identical. The Q_2/P_1^{disc} pair of finite element spaces was applied as spatial discretization. The quadrilateral grid depicted in Fig. 6.5 was used as initial grid (level 0). Table 7.4 gives information about the number of degrees of freedom for different refinement levels.

The nonlinear system in each discrete time was solved until the Euclidean norm of the residual vector was less than 10^{-10}. For the evaluation of the drag and lift coefficient with (D.16) and (D.17), the same approach was applied as described in Example 6.36. The temporal derivative in (D.16) and (D.17) was approximated by the backward difference formula

$$\partial_t \boldsymbol{u}_{n+1}^h \approx \frac{\boldsymbol{u}_{n+1}^h - \boldsymbol{u}_n^h}{\Delta t_{n+1}}.$$

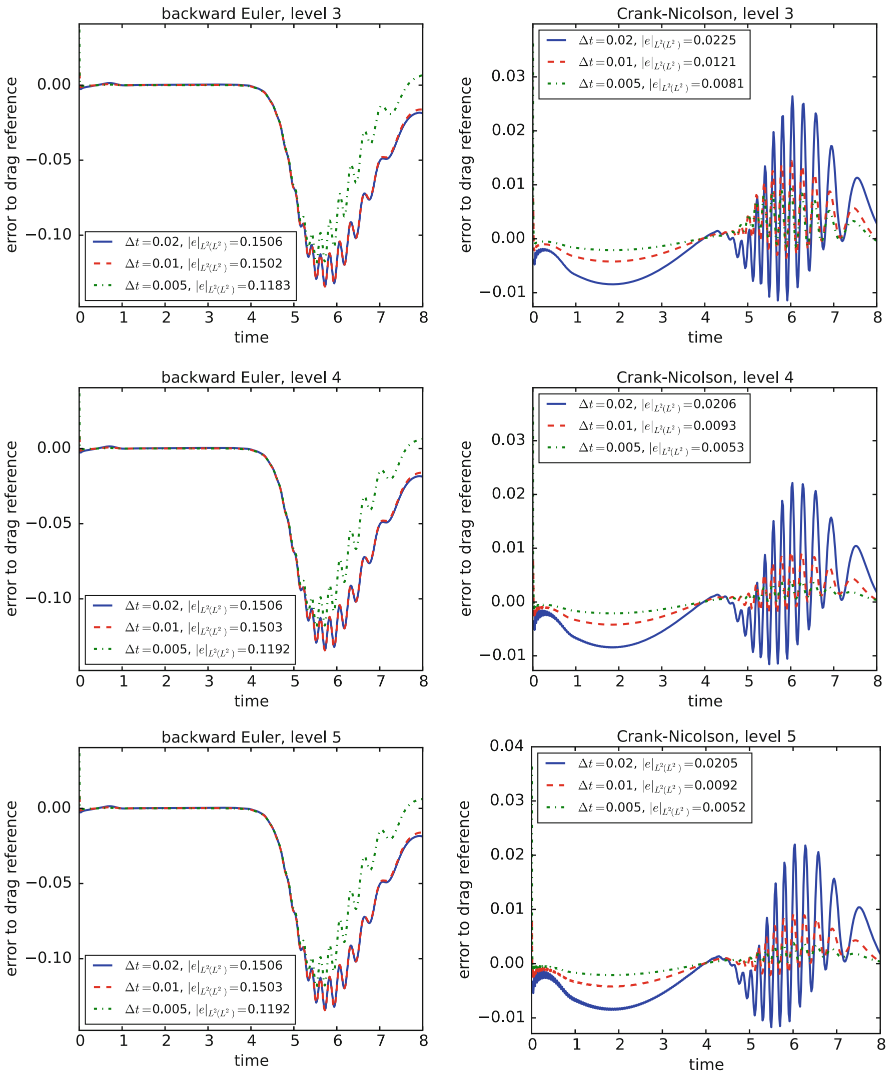

Fig. 7.3 Example 7.58. Error to the reference curve for the drag coefficient from Fig. D.10, Q_2/P_1^{disc}, levels 3–5 (*top to bottom*), *left*: backward Euler scheme, *right*: Crank–Nicolson, both for different lengths of the time step

Figures 7.3, 7.4, 7.5, 7.6, 7.7, and 7.8 present the obtained results. It can be observed that the use of the backward Euler scheme leads by far to the most inaccurate results among all time stepping schemes. In the considered example, the temporal and the spatial error are both of importance. Using the backward Euler scheme, the temporal error dominates the error in space. Considerably smaller time steps are necessary to reach a similar error level as obtained with the second order

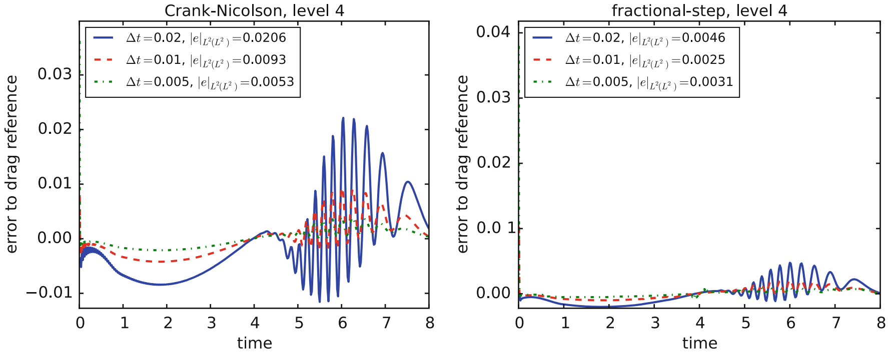

Fig. 7.4 Example 7.58. Error to the reference curve for the drag coefficient from Fig. D.10, Q_2/P_1^{disc}, *left*: Crank–Nicolson scheme, *right*: fractional-step θ-scheme

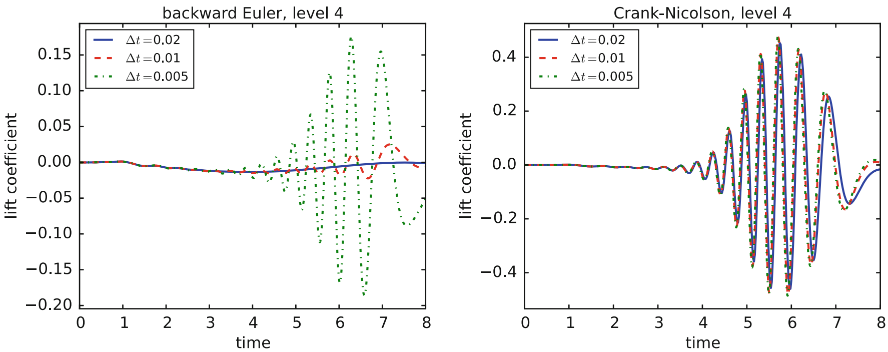

Fig. 7.5 Example 7.58. Temporal evolution of the lift coefficient, Q_2/P_1^{disc}, *left*: backward Euler scheme, *right*: Crank–Nicolson scheme. Note the different scales in both pictures

time stepping schemes. The temporal evolution of the lift coefficient, Fig. 7.5, shows that even the vortex shedding is not correctly predicted with the backward Euler scheme with time steps larger than or equal to $\Delta t = 0.005$. Altogether, this example demonstrates the qualitative difference of using first and second order time stepping schemes in combination with higher order discretizations in space for problems where the error in space does not dominate.

Also for the second order schemes, the error reductions with decreasing length of the time step can be clearly observed. Results with the second order time stepping schemes can be compared in Figs. 7.4, 7.7, and 7.8. Despite the same order of the Crank–Nicolson scheme and the fractional-step θ-scheme, it can be observed that the coefficients computed with the fractional-step θ-scheme are considerably more accurate. However, it will be discussed in Remark 7.59 that the numerical costs (computing times) for simulating the flow with the fractional-step θ-scheme are usually two to three times higher than the costs for the Crank–Nicolson scheme.

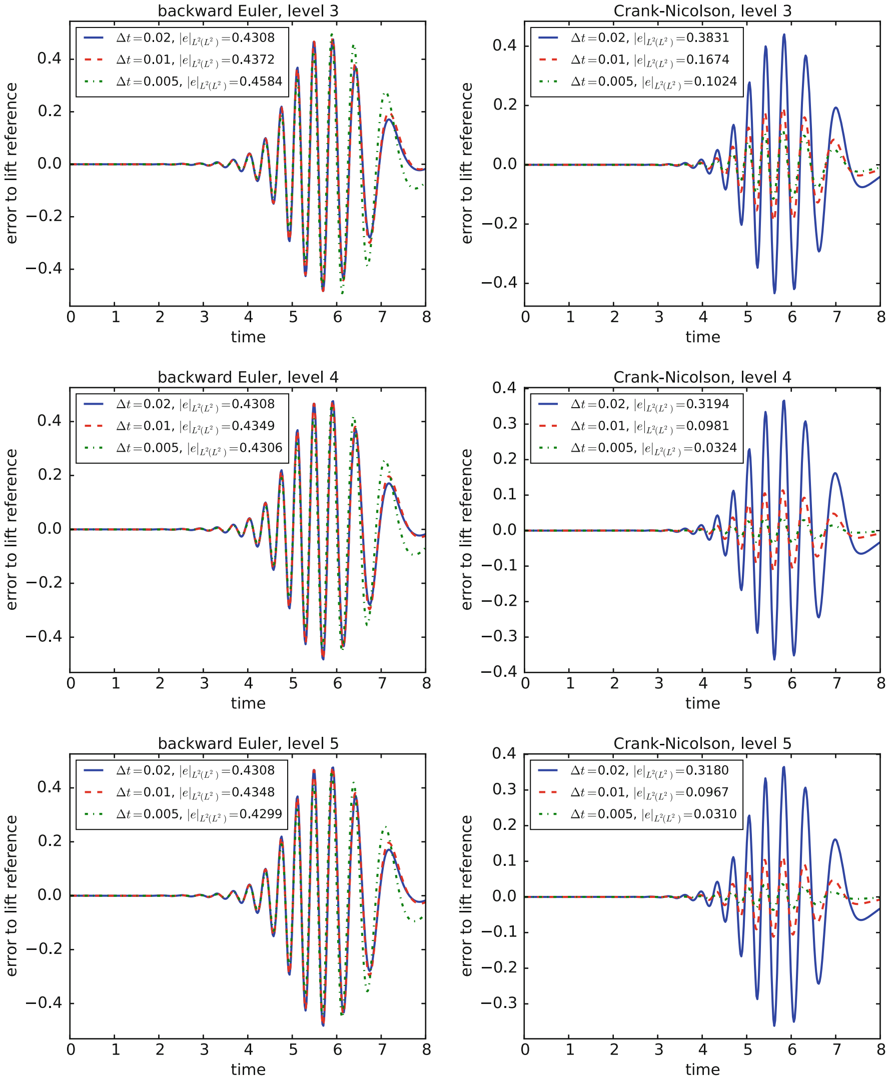

Fig. 7.6 Example 7.58. Error to the reference curve for the lift coefficient from Fig. D.10, Q_2/P_1^{disc}, levels 3–5 (*top to bottom*), *left*: backward Euler scheme, *right*: Crank–Nicolson, both for different lengths of the time step

The choice of the stopping criterion for solving the nonlinear problems in each discrete time is a delicate issue. If the stopping criterion is rather hard, the computation of the solution in the discrete times might become time-consuming and the overall computing time might become large. But one can expect to get accurate results. On the other hand, using a soft stopping criterion might speed up the simulations considerably. However, a rather inaccurate solution might be

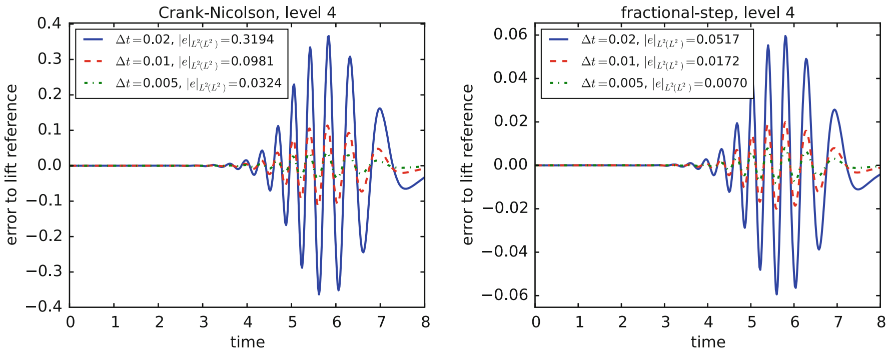

Fig. 7.7 Example 7.58. Error to the reference curve for the lift coefficient from Fig. D.10, Q_2/P_1^{disc}, *left*: Crank–Nicolson scheme, *right*: fractional-step θ-scheme. Note the different scales in the pictures

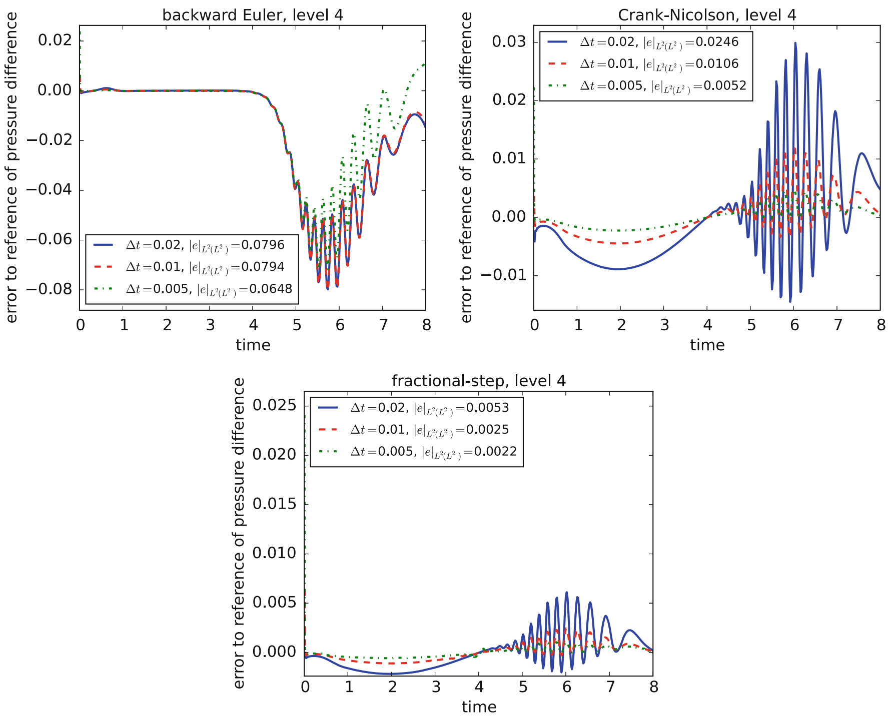

Fig. 7.8 Example 7.58. Error to the reference curve for the pressure difference between the front and the back of the cylinder from Fig. D.10, Q_2/P_1^{disc}, backward Euler scheme, Crank–Nicolson scheme, fractional-step θ-scheme (*left to right*, *top to bottom*)

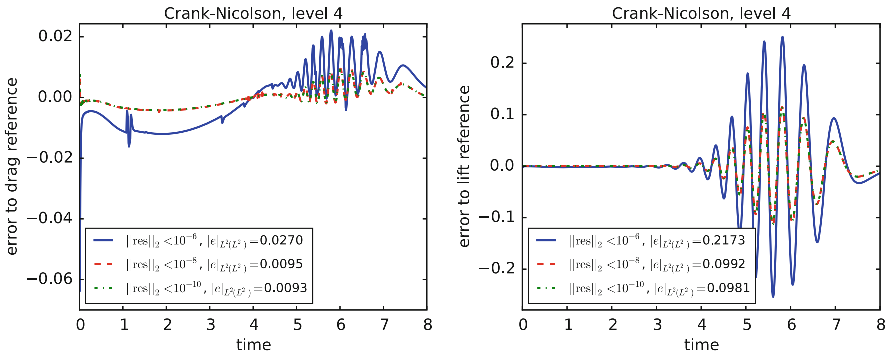

Fig. 7.9 Example 7.58. Dependency of the accuracy of the results on the stopping criterion for the Picard iteration in each discrete time, Q_2/P_1^{disc}, Crank–Nicolson scheme, $\Delta t = 0.01$

computed in some discrete times such that the overall error increases notably. This effect is demonstrated in Fig. 7.9. It can be seen that there is a notable gain in accuracy if as stopping criterion an Euclidean norm of the residual vector lower than or equal to 10^{-8} is used instead of 10^{-6}. It was found that the computing time for the tolerance 10^{-8} is around 2.5 times longer than for 10^{-6}. Decreasing the tolerance to 10^{-10} does not possess much impact on the accuracy compared with the tolerance 10^{-8}, however the computing time increased once more by a factor of around 2.8.

What does it mean 'hard' and 'soft' stopping criterion depends on the concrete example and on the used methods. □

Remark 7.59 (Numerical Experience with Implicit θ-Schemes for the Navier–Stokes Equations) The Crank–Nicolson and the fractional-step θ-scheme are already well tested and compared for the Navier–Stokes equations, see Emmrich (2001) for an overview. The Crank–Nicolson scheme is A-stable whereas the fractional-step θ-scheme is even strongly A-stable. That means, the Crank–Nicolson scheme may lead to numerical oscillations in problems with rough initial data or boundary conditions. These oscillations are damped out only if sufficiently small time steps are used. Compared with the fractional-step θ-scheme, a smaller time step might be necessary for the Crank–Nicolson scheme to ensure robustness. In numerical tests for the two-dimensional Navier–Stokes equations (John and Rang 2010; John et al. 2006a), the results obtained with the fractional-step θ-scheme were often somewhat more accurate than the results computed with the Crank–Nicolson scheme. However, the computing times for the fractional-step θ-scheme were in general two to three times longer than the computing times for the Crank–Nicolson scheme. Altogether, the ratio of accuracy per CPU time presented in the diagrams in John and Rang (2010) is somewhat better for the Crank–Nicolson scheme. □

Remark 7.60 (Implementation) The principal implementation of finite element methods for θ-schemes follows the same lines as the implementation for the Stokes and Oseen equations, compare Sect. 4.3 and Remark 5.19. Only the mass matrix

$$(M)_{ij} = \left(\boldsymbol{\phi}_j^h, \boldsymbol{\phi}_i^h\right)_K, \ i,j = 1,\ldots,3N_v, \tag{7.87}$$

is needed for representing the discretization of the temporal derivative. Then, the matrix of the algebraic system (7.85) for the Picard iteration, using the gradient form of the viscous term and the convective, divergence, or skew-symmetric form of the convective term, looks as follows

$$\begin{pmatrix} M & 0 & 0 & 0 \\ 0 & M & 0 & 0 \\ 0 & 0 & M & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \Delta t_{n+1} \begin{pmatrix} \theta_1 A_{11}\left(\boldsymbol{u}_{k+1}^{(m)}\right) & 0 & 0 & \frac{\Delta t_{k+1}}{\Delta t_{n+1}} B_1^T \\ 0 & \theta_1 A_{11}\left(\boldsymbol{u}_{k+1}^{(m)}\right) & 0 & \frac{\Delta t_{k+1}}{\Delta t_{n+1}} B_2^T \\ 0 & 0 & \theta_1 A_{11}\left(\boldsymbol{u}_{k+1}^{(m)}\right) & \frac{\Delta t_{k+1}}{\Delta t_{n+1}} B_3^T \\ \frac{\Delta t_{k+1}}{\Delta t_{n+1}} B_1 & \frac{\Delta t_{k+1}}{\Delta t_{n+1}} B_2 & \frac{\Delta t_{k+1}}{\Delta t_{n+1}} B_3 & 0 \end{pmatrix}.$$

The term with θ_2 on the right-hand side of (7.85) has the form

$$-\theta_2 \Delta t_{n+1} \begin{pmatrix} A_{11}(\boldsymbol{u}_k) & 0 & 0 & 0 \\ 0 & A_{11}(\boldsymbol{u}_k) & 0 & 0 \\ 0 & 0 & A_{11}(\boldsymbol{u}_k) & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \underline{u_k} \\ \underline{0} \end{pmatrix}, \tag{7.88}$$

where $A_{11}(\boldsymbol{u}_k)$ is the final matrix that was assembled in the Picard iteration at the discrete time t_k. Hence, the computation of (7.88) requires just a matrix-vector multiplication. □

Remark 7.61 (Semi-implicit Methods, IMEX Schemes) A popular approach is the use of semi-implicit schemes, so called IMEX (implicit-explicit) schemes, that avoid the solution of a nonlinear problem at each discrete time. Thus, performing simulations with IMEX schemes is potentially less time-consuming than with fully implicit schemes.

In these schemes, the nonlinear term $(\boldsymbol{u}_{k+1} \cdot \nabla)\,\boldsymbol{u}_{k+1}$, e.g., appearing in (7.83), is replaced by $\left(\boldsymbol{u}_{\text{prev}} \cdot \nabla\right) \boldsymbol{u}_{k+1}$, where $\boldsymbol{u}_{\text{prev}}$ can be obtained from already computed solutions. The simplest way consists in using $\boldsymbol{u}_{\text{prev}} = \boldsymbol{u}_k$, but often a linear extrapolation of the previous two time steps is used, i.e.,

$$\boldsymbol{u}_{\text{prev}} = \frac{\Delta t_{k+1}}{\Delta t_k}\left(\boldsymbol{u}_k - \boldsymbol{u}_{k-1}\right) + \boldsymbol{u}_k. \tag{7.89}$$

This linear extrapolation gives for $\theta = 0.5$ a similar scheme to the stabilized extrapolated Crank–Nicolson scheme CNLE(stab) proposed and studied in Ingram

(2013a). For constant time step, the convective term in Ingram (2013a) has the form

$$\frac{1}{4}\Delta t\Big[\,(((2\boldsymbol{u}_n+\boldsymbol{u}_{n-1}-\boldsymbol{u}_{n-2})\cdot\nabla)\,\boldsymbol{u}_{n+1},\boldsymbol{v})$$
$$+\,(((2\boldsymbol{u}_n+\boldsymbol{u}_{n-1}-\boldsymbol{u}_{n-2})\cdot\nabla)\,\boldsymbol{u}_n,\boldsymbol{v})\,\Big], \tag{7.90}$$

whereas with the extrapolation (7.89), one obtains the convective term

$$\frac{1}{2}\Delta t\Big[\,(((2\boldsymbol{u}_n-\boldsymbol{u}_{n-1})\cdot\nabla)\,\boldsymbol{u}_{n+1},\boldsymbol{v})+(((2\boldsymbol{u}_{n-1}-\boldsymbol{u}_{n-2})\cdot\nabla)\,\boldsymbol{u}_n,\boldsymbol{v})\,\Big]. \tag{7.91}$$

The second term in (7.90) and (7.91) appears in the right-hand side of the equation at time instance t_{n+1}. Note that

$$\frac{2}{4}\,(2\boldsymbol{u}_n+\boldsymbol{u}_{n-1}-\boldsymbol{u}_{n-2})=\frac{1}{2}\,(2\boldsymbol{u}_n-\boldsymbol{u}_{n-1})+\frac{1}{2}\,(2\boldsymbol{u}_{n-1}-\boldsymbol{u}_{n-2})\,,$$

such that (7.90) and (7.91) use a different splitting of the extrapolated velocity field to the terms at the discrete times t_{n+1} and t_n. The form (7.91) enables the computation of the right-hand side of the equation at t_{n+1} by using the matrix that was assembled at t_n, compare Remark 7.60. □

Example 7.62 (IMEX Schemes) To demonstrate the performance of IMEX schemes, the two-dimensional flow around a cylinder with steady-state inflow, Example D.8, will be considered. Simulations were performed for the Q_2/P_1^{disc} pair of finite element spaces, see Fig. 6.5 for the initial grid and Table 7.4 for information on the number of degrees of freedom. The Crank–Nicolson scheme with $\Delta t = 0.005$ was applied. The IMEX scheme was used with constant extrapolation $\boldsymbol{u}_{\mathrm{prev}} = \boldsymbol{u}_n$ and with the linear extrapolation (7.91).

Results obtained on level 4 are presented in Fig. 7.10. The results on level 3 and level 5 are qualitatively the same. It can be seen that the results for the IMEX scheme with constant extrapolation are quite inaccurate, e.g., the drag coefficient and the maximal value of the lift coefficient are considerably too large. The same behavior was observed in Caiazzo et al. (2014) for the backward Euler scheme with constant extrapolation. In contrast, using the linear extrapolation in the IMEX scheme gives almost the same accurate results as computed with the fully implicit scheme.

A numerical analysis for the drag and the lift coefficient computed with the backward Euler scheme with constant extrapolation can be found in Tabata and Tagami (2000). This analysis assumes sufficient regularity of the solution of the Navier–Stokes equations. Since domains for flows around bodies, where drag and lift are of importance, are usually not convex, it is not clear whether this assumption is satisfied.

The gain in efficiency depends on many aspects, like the concrete problem, the used solvers, details of the implementation, and others. For the numerical results presented in this example, the sparse direct solver umfpack, see Davis (2004) and

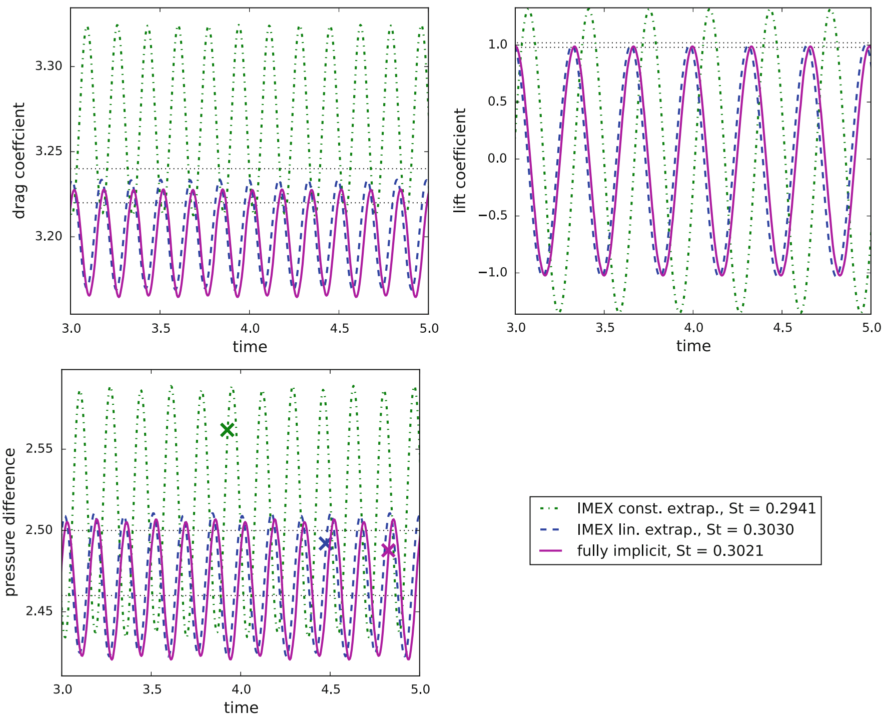

Fig. 7.10 Example 7.62. Quantities of interest. Reference intervals, see Table D.2, are depicted with the *dotted lines*. The *cross marks* the middle of the period. The reference interval for the Strouhal number is $[0.295, 0.305]$

Remark 9.5, was applied in the IMEX schemes. In the fully implicit scheme, a Picard iteration with inexact solutions of the linear systems of equations was used. The Euclidean norm of the residual was reduced by the factor 10 and the stopping criterion of the Picard iteration was that this norm should be smaller than 10^{-8}. As solver, FGMRES with a coupled multigrid preconditioner, see Sect. 9.2.2, was applied. The simulations with the IMEX scheme were faster, by a factor of around 40 on level 3, around 10 on level 4, and around 2.25 on level 5. The reason for the smaller gain on finer levels is that the iterative solver scales better than the direct sparse solver if the mesh is refined. □

Remark 7.63 (Adaptive Time Step Control) There are well understood techniques for an adaptive time step control in the numerical simulation of ordinary differential equations, e.g., compare Hairer et al. (1993, Sect. II.4). These techniques rely on comparing two solutions obtained with methods of different order and on procedures, so-called controllers, which propose the length of the next time step. If possible, so-called imbedded schemes are used, as an inexpensive approach, for computing a solution with one order less than for the original time stepping scheme.

However, there are no imbedded schemes for the class of θ-schemes such that other approaches for an adaptive time step control are necessary.

The approach proposed in Turek (1999) compares the results of the fractional-step θ-scheme and the Crank–Nicolson scheme. These schemes have a different constant in the leading term of their error expansions. This difference, together with the difference of the results obtained with both schemes, can be used to predict an appropriate length of the time step. The main drawback is the high computational effort of this approach. The step with the Crank-Nicolson scheme is used only to determine the size of the next time step. The costs of this step are not negligible such that the adaptive time step control increases the costs per time step notably. A similar but less expensive approach is proposed in Kay et al. (2010). In this approach, the step length control is based on comparing the (velocity) solutions of the Crank–Nicolson scheme and the explicit Adams–Bashforth method of second order. A more heuristic approach monitors just the change of the computed solution in some norm and the length of the time step is varied due to this change, e.g., see Berrone and Marro (2009). □

7.3.2 Other Schemes

Remark 7.64 (BDF2) The backward difference formula of order two, called BDF2, applied to the Navier–Stokes equations (7.1) has the form

$$
\begin{aligned}
&\frac{3}{2}\boldsymbol{u}_{n+1} + \Delta t_{n+1}\big[-\nu\Delta\boldsymbol{u}_{n+1} + (\boldsymbol{u}_{n+1}\cdot\nabla)\,\boldsymbol{u}_{n+1}\big] + \Delta t_{n+1}\nabla p_{n+1}\\
&\quad = 2\boldsymbol{u}_n - \frac{1}{2}\boldsymbol{u}_{n-1} - \Delta t_{n+1}\big[-\nu\Delta\boldsymbol{u}_n + (\boldsymbol{u}_n\cdot\nabla)\,\boldsymbol{u}_n\big] + \Delta t_{n+1}\boldsymbol{f}_{n+1},\\
&0 = \Delta t_{n+1}\nabla\cdot\boldsymbol{u}_{n+1},
\end{aligned}
\tag{7.92}
$$

for $n \geq 2$. The BDF2 scheme is of second order and it is strongly A-stable. Thus, compared with the Crank–Nicolson scheme, see Example 7.49, it is of the same order, more stable, but it requires the storage of $\boldsymbol{u}_{n-1}$. Because of its higher order and of its stability, BDF2 is a popular scheme, e.g., it is the standard temporal discretization in the software package deal.II, see Bangerth et al. (2007). □

Remark 7.65 (Higher Order Methods) Rosenbrock methods of order 3 with $s = 3$ or $s = 4$ stages were compared with θ-schemes in John et al. (2006a). These methods are linearly implicit Runge–Kutta schemes. In each discrete time, s linear saddle point problems with the same system matrix have to be solved. Rosenbrock methods allow an efficient time step control with imbedded schemes. The numerical solutions for the two-dimensional problems studied in John et al. (2006a) were generally considerably more accurate for the Rosenbrock methods compared with the θ-schemes. However, the simulations with the θ-schemes were clearly more efficient. Rosenbrock methods up to fourth order and diagonally implicit Runge–Kutta methods (DIRK methods) were studied and compared with θ-schemes in John and Rang (2010) at a two-dimensional flow around a cylinder, see Example D.9. All higher order schemes, i.e., all schemes that are at least of third order, allow an adaptive time step control via imbedded methods. If the time step control was applied, then many of the higher order methods outperformed the θ-schemes clearly with respect to the ratio of accuracy and CPU time.

Nevertheless, the use of higher order time stepping schemes does not seem to be popular at the moment. □

7.4 Finite Element Error Analysis: The Fully Discrete Case

Remark 7.66 (Comparison to the Continuous-in-Time Case and Contents of this Section) The numerical analysis of the fully discretized Navier–Stokes equations is similar to the continuous-in-time case presented in Sect. 7.2. However, the discretization of the temporal derivative introduces a consistency error, whose estimate will be presented in Lemma 7.67. In the fully discrete case, sums over the discrete times appear in the formulas instead of time-space norms. Thus, the presentations are less compact and somewhat more involved compared with the continuous-in-time case.

This section presents the analysis for the fully discrete case at the example of the backward Euler discretization in time in combination with the Galerkin finite element method in space. For brevity, the presentation of the error analysis is restricted to estimates for the velocity. □

Lemma 7.67 (Consistency Errors for Discretizations of Temporal Derivatives) *Let* $v, \partial_t v, \partial_{tt} v \in L^2\left(t_n, t_{n+1}; L^2(\Omega)\right)$*, then*

$$\left\| \partial_t v_{n+1} - \frac{v_{n+1} - v_n}{\Delta t} \right\|_{L^2(\Omega)}^2 \le \Delta t \left\| \partial_{tt} v \right\|_{L^2\left(t_n, t_{n+1}; L^2(\Omega)\right)}^2 . \tag{7.93}$$

If in addition $\partial_{ttt} v \in L^2\left(t_n, t_{n+1}; L^2(\Omega)\right)$*, then*

$$\left\| \partial_t v_{n+1/2} - \frac{v_{n+1} - v_n}{\Delta t} \right\|_{L^2(\Omega)}^2 \le \Delta t^3 \left\| \partial_{ttt} v \right\|_{L^2\left(t_n, t_{n+1}; L^2(\Omega)\right)}^2 . \tag{7.94}$$

Proof Taylor's formula with integral remainder or, likewise, integration by parts gives

$$v_n = v_{n+1} - \Delta t \partial_t v_{n+1} + \int_{t_n}^{t_{n+1}} (t - t_n)\, \partial_{tt} v \; dt,$$

which is equivalent to

$$\partial_t v_{n+1} - \frac{v_{n+1} - v_n}{\Delta t} = \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} (t - t_n)\, \partial_{tt} v \, dt.$$

Then, one obtains with the Cauchy–Schwarz inequality (A.10), observing that one of the arising factors does not depend on space, and the Theorem of Fubini

$$\begin{aligned}
&\left\| \partial_t v_{n+1} - \frac{v_{n+1} - v_n}{\Delta t} \right\|_{L^2(\Omega)}^2 \\
&\le \frac{1}{\Delta t^2} \left\| \left(\int_{t_n}^{t_{n+1}} (t - t_n)^2 \, dt \right)^{1/2} \left(\int_{t_n}^{t_{n+1}} (\partial_{tt} v)^2 \, dt \right)^{1/2} \right\|_{L^2(\Omega)}^2 \\
&= \frac{1}{\Delta t^2} \left(\int_{t_n}^{t_{n+1}} (t - t_n)^2 \, dt \right) \left(\int_\Omega \int_{t_n}^{t_{n+1}} (\partial_{tt} v)^2 \, dt d\boldsymbol{x} \right) \\
&= \frac{1}{\Delta t^2} \frac{\Delta t^3}{3} \int_{t_n}^{t_{n+1}} \int_\Omega (\partial_{tt} v)^2 \, d\boldsymbol{x} dt,
\end{aligned}$$

such that (7.93) follows. Note that for the sake of simplicity, the factor $1/3$ is estimated by 1.

The proof of (7.94) proceeds the same way. A Taylor series expansion or successive integration by parts gives for $\theta \in [0, 1]$

$$\begin{aligned}
v_n &= v_{n+\theta} - \theta \Delta t \partial_t v_{n+\theta} + \frac{\theta^2}{2} \Delta t^2 \partial_{tt} v_{n+\theta} + \frac{1}{2} \int_{t_{n+\theta}}^{t_n} (t_n - t)^2 \partial_{ttt} v \, dt, \\
v_{n+1} &= v_{n+\theta} + (1 - \theta) \Delta t \partial_t v_{n+\theta} + \frac{(1 - \theta)^2}{2} \Delta t^2 \partial_{tt} v_{n+\theta} + \frac{1}{2} \int_{t_{n+\theta}}^{t_{n+1}} (t_{n+1} - t)^2 \partial_{ttt} v \, dt.
\end{aligned}$$

Subtracting the second equation from the first equation and rearranging terms yields

$$\begin{aligned}
\partial_t v_{n+\theta} = &\frac{v_{n+1} - v_n}{\Delta t} + \frac{\Delta t}{2} \left(\theta^2 - (1 - \theta)^2 \right) \partial_{tt} v_{n+\theta} \\
&- \frac{1}{2\Delta t} \int_{t_n}^{t_{n+\theta}} (t_n - t)^2 \partial_{ttt} v \, dt - \frac{1}{2\,\Delta t} \int_{t_{n+\theta}}^{t_{n+1}} (t_{n+1} - t)^2 \partial_{ttt} v \, dt.
\end{aligned}$$

The term with $\partial_{tt} v_{n+\theta}$ vanishes only if $\theta = 1/2$. Considering $\theta = 1/2$, one obtains in the same way as in the proof of (7.93)

$$\begin{aligned}
&\left\| \partial_t v_{n+1/2} - \frac{v_{n+1} - v_n}{\Delta t} \right\|_{L^2(\Omega)}^2 \\
&\le \frac{1}{4\Delta t^2} \left[\left(\int_{t_n}^{t_{n+1/2}} (t_n - t)^4 \, dt \right) \left(\int_\Omega \int_{t_n}^{t_{n+1/2}} (\partial_{ttt} v)^2 \, dt d\boldsymbol{x} \right) \right. \\
&\quad \left. + \left(\int_{t_{n+1/2}}^{t_{n+1}} (t_{n+1} - t)^4 \, dt \right) \left(\int_\Omega \int_{t_{n+1/2}}^{t_{n+1}} (\partial_{ttt} v)^2 \, dt d\boldsymbol{x} \right) \right],
\end{aligned}$$

from which (7.94) is obtained with a straightforward calculation. ■

Remark 7.68 (The Backward Euler Scheme) The backward Euler scheme with equidistant time step has the following form: Given $\boldsymbol{f}_{n+1} \in V'$ and $\boldsymbol{u}_n^h \in V^h$, compute $\left(\boldsymbol{u}_{n+1}^h, p_{n+1}^h\right) \in V^h \times Q^h$, $n = 0, 1, 2, \ldots$ by solving,

$$\begin{aligned}
&\frac{1}{\Delta t} \left(\boldsymbol{u}_{n+1}^h - \boldsymbol{u}_n^h, \boldsymbol{v}^h\right) + \nu \left(\nabla \boldsymbol{u}_{n+1}^h, \nabla \boldsymbol{v}^h\right) + n_{\text{skew}} \left(\boldsymbol{u}_{n+1}^h, \boldsymbol{u}_{n+1}^h, \boldsymbol{v}^h\right) \qquad (7.95) \\
&- \left(\nabla \cdot \boldsymbol{v}^h, p_{n+1}^h\right) + \left(\nabla \cdot \boldsymbol{u}_{n+1}^h, q^h\right) = \langle \boldsymbol{f}_{n+1}, \boldsymbol{v}^h \rangle_{V',V} \quad \forall \left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h.
\end{aligned}$$

The initial condition $\boldsymbol{u}_0^h$ is some interpolation or projection of $\boldsymbol{u}_0$ into the finite element space, e.g., the Lagrange interpolation or the Stokes projection (4.54). □

Lemma 7.69 (Existence, Uniqueness, and Stability of the Finite Element Solution) *Let V^h and Q^h be a pair of finite element spaces that satisfy the discrete inf-sup condition* (3.51)*. Problem* (7.95) *possesses a unique solution for sufficiently small time steps. The time step restriction depends on the mesh width.*

It holds for any solution of (7.95) *and for all $N \ge 0$ that*

$$\left\| \boldsymbol{u}_{N+1}^h \right\|_{L^2(\Omega)}^2 + \nu \Delta t \sum_{n=0}^N \left\| \nabla \boldsymbol{u}_{n+1}^h \right\|_{L^2(\Omega)}^2 \le \left\| \boldsymbol{u}_0^h \right\|_{L^2(\Omega)}^2 + \frac{\Delta t}{\nu} \sum_{n=0}^N \left\| \boldsymbol{f}_{n+1} \right\|_{H^{-1}(\Omega)}^2 . \qquad (7.96)$$

Proof First, the stability bound (7.96) will be proved since the result will be used for showing existence and uniqueness of a solution.

Stability The stability proof proceeds in the usual way by taking any solution of (7.95) as test function $\left(\boldsymbol{v}^h, q^h\right) = \left(\boldsymbol{u}_{n+1}^h, p_{n+1}^h\right)$. Using the skew-symmetry (6.26) of the convective term, which gives $n_{\text{skew}} \left(\boldsymbol{u}_{n+1}^h, \boldsymbol{u}_{n+1}^h, \boldsymbol{u}_{n+1}^h\right) = 0$, leads to

$$\frac{1}{\Delta t} \left(\boldsymbol{u}_{n+1}^h - \boldsymbol{u}_n^h, \boldsymbol{u}_{n+1}^h\right) + \nu \left\| \nabla \boldsymbol{u}_{n+1}^h \right\|_{L^2(\Omega)}^2 = \langle \boldsymbol{f}_{n+1}, \boldsymbol{u}_{n+1}^h \rangle_{V',V}. \qquad (7.97)$$

Analogously to the algebraic relation $a^2 - ab = \left(a^2 + (a-b)^2 - b^2\right)/2$ for real numbers a and b, it holds

$$\left(\boldsymbol{u}_{n+1}^h - \boldsymbol{u}_n^h, \boldsymbol{u}_{n+1}^h\right) = \frac{1}{2}\left(\left\|\boldsymbol{u}_{n+1}^h\right\|_{L^2(\Omega)}^2 + \left\|\boldsymbol{u}_{n+1}^h - \boldsymbol{u}_n^h\right\|_{L^2(\Omega)}^2 - \left\|\boldsymbol{u}_n^h\right\|_{L^2(\Omega)}^2\right). \tag{7.98}$$

Inserting (7.98) in (7.97) and using that $\left\|\boldsymbol{u}_{n+1}^h - \boldsymbol{u}_n^h\right\|_{L^2(\Omega)}^2 \geq 0$ gives the estimate

$$\left\|\boldsymbol{u}_{n+1}^h\right\|_{L^2(\Omega)}^2 + 2\nu\Delta t\left\|\nabla\boldsymbol{u}_{n+1}^h\right\|_{L^2(\Omega)}^2 \leq \left\|\boldsymbol{u}_n^h\right\|_{L^2(\Omega)}^2 + 2\Delta t\langle\boldsymbol{f}_{n+1}, \boldsymbol{u}_{n+1}^h\rangle_{V',V}.$$

Applying the estimate of the dual pairing and Young's inequality (A.5) yields

$$\left\|\boldsymbol{u}_{n+1}^h\right\|_{L^2(\Omega)}^2 + \nu\Delta t\left\|\nabla\boldsymbol{u}_{n+1}^h\right\|_{L^2(\Omega)}^2 \leq \left\|\boldsymbol{u}_n^h\right\|_{L^2(\Omega)}^2 + \frac{\Delta t}{\nu}\left\|\boldsymbol{f}_{n+1}\right\|_{H^{-1}(\Omega)}^2.$$

Now, the stability estimate (7.96) is obtained by taking the sum from 0 to N and observing that the first terms on both sides constitute a telescopic sum.

Existence and Uniqueness of a Solution In each discrete time of the backward Euler scheme (7.95), a nonlinear discrete problem has to be solved which is similar to the steady-state Navier–Stokes equations. Only the term arising from the temporal discretization appears additionally. The proof of the existence and uniqueness of a solution will be performed in the same way as for the steady-state Navier–Stokes equations, compare the proof of Theorem 6.17. Consider the problem of Oseen type: Given $\boldsymbol{f}_{n+1} \in V'$ and $\boldsymbol{b}^h, \boldsymbol{u}_n^h \in V_{\text{div}}^h$, compute $\boldsymbol{u}^h \in V_{\text{div}}^h$, by solving,

$$\frac{1}{\Delta t}\left(\boldsymbol{u}^h, \boldsymbol{v}^h\right) + \nu\left(\nabla\boldsymbol{u}^h, \nabla\boldsymbol{v}^h\right) + n_{\text{skew}}\left(\boldsymbol{b}^h, \boldsymbol{u}^h, \boldsymbol{v}^h\right) = \frac{1}{\Delta t}\left(\boldsymbol{u}_n^h, \boldsymbol{v}^h\right) + \langle\boldsymbol{f}_{n+1}, \boldsymbol{v}^h\rangle_{V',V} \tag{7.99}$$

for all $\boldsymbol{v}^h \in V_{\text{div}}^h$. With this problem, the following operator is defined

$$N_{\text{conv}}^h : V_{\text{div}}^h \to V_{\text{div}}^h, \quad \boldsymbol{b}^h \mapsto \boldsymbol{u}^h.$$

Each fixed point of N_{conv}^h is a velocity solution of the backward Euler problem (7.95) at time t_{n+1}.

The existence and uniqueness of a solution of (7.99) can be proved with the Theorem of Lax–Milgram, see the proof of Lemma 7.77 for details. Since the convective term in (7.99) vanishes for $\boldsymbol{v}^h = \boldsymbol{u}^h$, one gets analogously to the first part of the proof of the current lemma the stability estimate

$$\left\|\boldsymbol{u}^h\right\|_{L^2(\Omega)}^2 + \nu\Delta t\left\|\nabla\boldsymbol{u}^h\right\|_{L^2(\Omega)}^2 \leq \left\|\boldsymbol{u}_n^h\right\|_{L^2(\Omega)}^2 + \frac{\Delta t}{\nu}\left\|\boldsymbol{f}_{n+1}\right\|_{H^{-1}(\Omega)}^2. \tag{7.100}$$

Neglecting the first term on the left-hand side, it follows that N^h_{conv} is bounded independently of $\boldsymbol{b}^h$

$$\begin{aligned}\left\| N^h_{\text{conv}} \right\| &= \sup_{\boldsymbol{b}^h \in V^h_{\text{div}}, \left\| \boldsymbol{b}^h \right\|_V = 1} \left\| N^h_{\text{conv}} \boldsymbol{b}^h \right\|_V = \sup_{\boldsymbol{b}^h \in V^h_{\text{div}}, \left\| \boldsymbol{b}^h \right\|_V = 1} \left\| \boldsymbol{u}^h \right\|_V \\ &\le \left(\frac{1}{\nu \Delta t} \left\| \boldsymbol{u}^h_n \right\|^2_{L^2(\Omega)} + \frac{1}{\nu^2} \left\| \boldsymbol{f}_{n+1} \right\|^2_{H^{-1}(\Omega)} \right)^{1/2}.\end{aligned}$$

With the inverse inequality (C.37) applied to the first term on the left-hand side of (7.100), one obtains

$$\left\| \nabla \boldsymbol{u}^h \right\|^2_{L^2(\Omega)} \le \frac{1}{C^{-2}_{\text{inv}} h^2 + \nu \Delta t} \left(\left\| \boldsymbol{u}^h_n \right\|^2_{L^2(\Omega)} + \frac{\Delta t}{\nu} \left\| \boldsymbol{f}_{n+1} \right\|^2_{H^{-1}(\Omega)} \right). \tag{7.101}$$

Choosing $\boldsymbol{b}^h_1, \boldsymbol{b}^h_2 \in V^h_{\text{div}}$ arbitrarily, setting $\boldsymbol{u}^h_1 = N^h_{\text{conv}} \boldsymbol{b}^h_1$, $\boldsymbol{u}^h_2 = N^h_{\text{conv}} \boldsymbol{b}^h_2$, and subtracting the corresponding problems yields for $\boldsymbol{v}^h = \boldsymbol{u}^h_1 - \boldsymbol{u}^h_2$

$$\frac{1}{\Delta t} \left\| \boldsymbol{u}^h_1 - \boldsymbol{u}^h_2 \right\|^2_{L^2(\Omega)} + \nu \left\| \nabla \left(\boldsymbol{u}^h_1 - \boldsymbol{u}^h_2 \right) \right\|^2_{L^2(\Omega)} = -n_{\text{skew}} \left(\boldsymbol{b}^h_1 - \boldsymbol{b}^h_2, \boldsymbol{u}^h_1, \boldsymbol{u}^h_1 - \boldsymbol{u}^h_2 \right). \tag{7.102}$$

Applying the inverse inequality to the first term on the left-hand side in (7.102) and estimate (6.29) of the convective term gives

$$\left(\frac{C^{-2}_{\text{inv}} h^2}{\Delta t} + \nu \right) \left\| \nabla \left(\boldsymbol{u}^h_1 - \boldsymbol{u}^h_2 \right) \right\|^2_{L^2(\Omega)} \le N^h_{\text{div}} \left\| \boldsymbol{b}^h_1 - \boldsymbol{b}^h_2 \right\|_V \left\| \boldsymbol{u}^h_1 \right\|_V \left\| \nabla \left(\boldsymbol{u}^h_1 - \boldsymbol{u}^h_2 \right) \right\|_{L^2(\Omega)},$$

with N^h_{div} defined in (6.57), such that, inserting (7.101),

$$\begin{aligned}&\left\| \nabla \left(\boldsymbol{u}^h_1 - \boldsymbol{u}^h_2 \right) \right\|_{L^2(\Omega)} \\ &\le N^h_{\text{div}} \Delta t \left(\frac{1}{C^{-2}_{\text{inv}} h^2 + \nu \Delta t} \right)^{3/2} \left(\left\| \boldsymbol{u}^h_n \right\|^2_{L^2(\Omega)} + \frac{\Delta t}{\nu} \left\| \boldsymbol{f}_{n+1} \right\|^2_{H^{-1}(\Omega)} \right)^{1/2} \left\| \boldsymbol{b}^h_1 - \boldsymbol{b}^h_2 \right\|_V.\end{aligned}$$

For sufficiently small time steps, the factor in front of $\left\| \boldsymbol{b}^h_1 - \boldsymbol{b}^h_2 \right\|_V$ becomes smaller than 1. In this situation, the existence and uniqueness of a velocity solution of (7.95) follows from the fixed point theorem of Banach, see Theorem A.68. The existence and uniqueness of the pressure at t_{n+1} is a consequence of the satisfaction of the discrete inf-sup condition. ■

Remark 7.70 (To the Existence and Uniqueness of a Solution) Problem 7.95 is of similar form as the steady-state Navier–Stokes equations. The existence of a solution for an arbitrary length of the time step can be proved in the same way as the existence of the solution of the steady-state Navier–Stokes equations, compare Marion and Temam (1998, Sect. 15.2) and the references in the proof of Theorem 6.17.

The uniqueness of a solution for sufficiently small time steps without dependency on the mesh size can be proved with higher regularity assumptions on the solution. For instance, based on the assumption $\nabla \boldsymbol{u} \in L^\infty\left((0,T);L^2(\Omega)\right)$, it was shown in Heywood and Rannacher (1990, (3.15)) that $\left\|\nabla \boldsymbol{u}^h\right\|_{L^\infty\left((0,T);L^2(\Omega)\right)} < C$ holds independently of h. Using this bound in (7.101) gives a condition on the time step that does not depend on the mesh width.

Alternatively, having proved the existence of a solution, then the uniqueness for small time steps can be proved by assuming that there are two solutions $\boldsymbol{u}_1^h$ and $\boldsymbol{u}_2^h$, subtracting the equations for both of them, and using as test function the difference

$$
\begin{aligned}
&\frac{1}{\Delta t}\left\|\boldsymbol{u}_1^h-\boldsymbol{u}_2^h\right\|_{L^2(\Omega)}^2+\nu\left\|\nabla\left(\boldsymbol{u}_1^h-\boldsymbol{u}_2^h\right)\right\|_{L^2(\Omega)}^2\\
&=-n_{\text{skew}}\left(\boldsymbol{u}_1^h,\boldsymbol{u}_1^h,\boldsymbol{u}_1^h-\boldsymbol{u}_2^h\right)+n_{\text{skew}}\left(\boldsymbol{u}_2^h,\boldsymbol{u}_1^h,\boldsymbol{u}_1^h-\boldsymbol{u}_2^h\right) \qquad (7.103)\\
&=n_{\text{conv}}\left(\boldsymbol{u}_1^h-\boldsymbol{u}_2^h,\boldsymbol{u}_1^h,\boldsymbol{u}_1^h-\boldsymbol{u}_2^h\right)+\frac{1}{2}\left(\nabla\cdot\left(\boldsymbol{u}_1^h-\boldsymbol{u}_2^h\right),\boldsymbol{u}_1^h\cdot\left(\boldsymbol{u}_1^h-\boldsymbol{u}_2^h\right)\right),
\end{aligned}
$$

where the definition of the skew-symmetric form (6.25) and (6.13) were used. Assuming $\boldsymbol{u}_1^h \in L^\infty\left((0,T);L^\infty(\Omega)\right)$, the right-hand side of (7.103) can be estimated by

$$
\frac{3}{2}\left\|\boldsymbol{u}_1^h\right\|_{L^\infty(\Omega)}\left\|\boldsymbol{u}_1^h-\boldsymbol{u}_2^h\right\|_{L^2(\Omega)}\left\|\nabla\left(\boldsymbol{u}_1^h-\boldsymbol{u}_2^h\right)\right\|_{L^2(\Omega)}.
$$

With this assumption, $\left\|\boldsymbol{u}_1^h\right\|_{L^\infty(\Omega)}$ can be bounded independently of h, see de Frutos et al. (2008, (2.16) and Remark 5.1). Then, the application of Young's inequality (A.5) leads to an inequality that can be only satisfied if the time step is sufficiently small, but without dependency on h. □

Remark 7.71 (Approximation of Time-Space Norms) Let X be a Banach space with norm $\|\cdot\|_X$. Then, the pth power of the norm of a function $v \in L^p(t_0,t_N;X)$, $p \in [1,\infty)$, $0 \le t_0 < t_N$, is given by

$$
\|v\|_{L^p(t_0,t_N;X)}^p=\int_{t_0}^{t_N}\|v\|_X^p\,d\tau. \qquad (7.104)
$$

Formula (7.104) can be approximated with the composite trapezoidal rule. To this end, the interval $[t_0,t_N]$ is decomposed in equidistant subintervals of length Δt. In each subinterval, the trapezoidal rule

$$
\int_{t_n}^{t_{n+1}}\|v\|_X^p\,d\tau \approx \frac{\Delta t}{2}\left(\|v(t_{n+1})\|_X^p+\|v(t_n)\|_X^p\right)
$$

is applied. The sum over all subintervals gives the following approximation of (7.104)

$$\|v\|^p_{L^p(t_0,t_N;X)} \approx \Delta t \left(\frac{\|v(t_0)\|^p_X}{2} + \sum_{n=1}^{N-1} \|v(t_n)\|^p_X + \frac{\|v(t_N)\|^p_X}{2} \right).$$

It follows that the terms with the sums in (7.96) are approximations of the composite trapezoidal rule, e.g., one has

$$\left\|\nabla \boldsymbol{u}^h\right\|^2_{L^2(t_1,t_{N+1};L^2(\Omega))} \approx \Delta t \sum_{n=0}^{N} \left\|\nabla \boldsymbol{u}^h_{n+1}\right\|^2_{L^2(\Omega)}.$$

□

Theorem 7.72 (Finite Element Error Estimate for the Velocity Computed with the Backward Euler Method in Combination with the Galerkin Finite Element Method) *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, be a bounded domain with polyhedral and Lipschitz continuous boundary, let $\boldsymbol{f} \in L^2(0,T;V')$, $\boldsymbol{u}_0 \in H_{\mathrm{div}}(\Omega)$, $\boldsymbol{u}_0^h \in V^h_{\mathrm{div}}$, and assume the following regularities for the solution $(\boldsymbol{u},p)$ of the Navier–Stokes equations* (7.40)

$$\begin{aligned} &\partial_{tt}\boldsymbol{u} \in L^2\left(0,T;H^1(\Omega)\right), \quad \partial_t\boldsymbol{u} \in L^2\left(0,T;V'\right), \\ &\nabla\boldsymbol{u} \in L^4\left(0,T;L^2(\Omega)\right), \quad p \in L^2\left(0,T;L^2(\Omega)\right). \end{aligned} \tag{7.105}$$

Consider the discretization with the backward Euler method in time and the Galerkin finite element method in space, see (7.95)*, with an inf-sup stable pair of spaces $V^h \times Q^h$ and with an equidistant time step Δt, where $t_n = n\Delta t$. Let the assumptions of Lemma 7.69 be satisfied and let the time step be sufficiently small such that*

$$\alpha_n \Delta t = C\frac{\|\nabla \boldsymbol{u}_n\|^4_{L^2(\Omega)}}{\nu^3}\Delta t < 1, \quad \forall\, n = 1,\ldots,N+1, \tag{7.106}$$

then the following error estimate holds for all $N \geq 0$

$$\begin{aligned} &\left\|\boldsymbol{u}_{N+1} - \boldsymbol{u}^h_{N+1}\right\|^2_{L^2(\Omega)} + \nu\Delta t \sum_{n=1}^{N+1} \left\|\nabla\left(\boldsymbol{u}_n - \boldsymbol{u}^h_n\right)\right\|^2_{L^2(\Omega)} \\ &\leq C\Bigg\{ \left\|\boldsymbol{u}_{N+1} - I^h_{\mathrm{St}}\boldsymbol{u}_{N+1}\right\|^2_{L^2(\Omega)} + \nu\Delta t \sum_{n=1}^{N+1} \left\|\nabla\left(\boldsymbol{u}_n - I^h_{\mathrm{St}}\boldsymbol{u}_n\right)\right\|^2_{L^2(\Omega)} \\ &\quad + \exp\left(\Delta t \sum_{n=1}^{N+1} \frac{\alpha_n}{1-\Delta t\alpha_n}\right)\Bigg[\left\|\boldsymbol{u}^h_0 - I^h_{\mathrm{St}}\boldsymbol{u}(0)\right\|^2_{L^2(\Omega)} \end{aligned}$$

$$
\begin{aligned}
&+\frac{(\Delta t)^2}{\nu}\left\|I_{\mathrm{St}}^h\left(\partial_{tt}\boldsymbol{u}\right)\right\|_{L^2\left(0,t_{N+1};L^2(\Omega)\right)}^2+\Delta t\sum_{n=1}^{N+1}\left(\frac{1}{\nu}\left(\left\|\partial_t\left(\boldsymbol{u}_n-I_{\mathrm{St}}^h\boldsymbol{u}_n\right)\right\|_{V'}^2\right.\right.\\
&\left.\left.+\left\|\boldsymbol{u}_n-I_{\mathrm{St}}^h\boldsymbol{u}_n\right\|_{L^2(\Omega)}\left\|\nabla\left(\boldsymbol{u}_n-I_{\mathrm{St}}^h\boldsymbol{u}_n\right)\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{u}_n\right\|_{L^2(\Omega)}^2\right)\right)\\
&+\inf_{q^h\in Q^h}\frac{\Delta t}{\nu}\sum_{n=1}^{N+1}\left\|p_n-q^h\right\|_{L^2(\Omega)}^2\\
&+\frac{1}{\nu^{3/2}}\left(\left\|\boldsymbol{u}_0^h\right\|_{L^2(\Omega)}^2+\frac{\Delta t}{\nu}\sum_{n=0}^{N}\left\|\boldsymbol{f}_{n+1}\right\|_{H^{-1}(\Omega)}^2\right) \qquad (7.107)\\
&\times\left(\Delta t\sum_{n=1}^{N+1}\left\|\nabla\left(\boldsymbol{u}_n-I_{\mathrm{St}}^h\boldsymbol{u}_n\right)\right\|_{L^2(\Omega)}^4\right)^{1/2}\Bigg]\Bigg\},
\end{aligned}
$$

with α_n defined in (7.106) and $I_{\mathrm{St}}^h\boldsymbol{u}_n$ being the Stokes projection of $\boldsymbol{u}_n$ with $p=0$, see (4.54), for which $\partial_t I_{\mathrm{St}}^h\boldsymbol{u}\in L^2(0,T;V')$ is assumed.

Proof The proof proceeds principally with the same steps as the proof of Theorem 7.35:

1. derivation of an error equation and splitting of the error,
2. estimate all terms on the right hand-side of the error equation,
3. application of a discrete Gronwall lemma,
4. application of the triangle inequality.

1. *Derivation of an error equation and splitting of the error.* The error at t_{n+1} is decomposed in an approximation error and a finite element remainder. The splitting is performed with the Stokes projection defined in (4.54) with $p=0$

$$
\begin{aligned}
\boldsymbol{e}(t_{n+1})&=\boldsymbol{u}(t_{n+1})-\boldsymbol{u}^h(t_{n+1})=\left(\boldsymbol{u}(t_{n+1})-I_{\mathrm{St}}^h\boldsymbol{u}(t_{n+1})\right)+\left(I_{\mathrm{St}}^h\boldsymbol{u}(t_{n+1})-\boldsymbol{u}^h(t_{n+1})\right)\\
&=\boldsymbol{\eta}(t_{n+1})-\boldsymbol{\phi}^h(t_{n+1})=\boldsymbol{\eta}_{n+1}-\boldsymbol{\phi}_{n+1}^h. \qquad (7.108)
\end{aligned}
$$

By construction, it is $\boldsymbol{\phi}_{n+1}^h\in V_{\mathrm{div}}^h$. The error equation is obtained by subtracting (7.95) from (7.40). Considering only test functions $\boldsymbol{v}^h\in V_{\mathrm{div}}^h$, one obtains

$$
\begin{aligned}
&\left(\partial_t\boldsymbol{u}_{n+1}-\frac{\boldsymbol{u}_{n+1}^h-\boldsymbol{u}_n^h}{\Delta t},\boldsymbol{v}^h\right)+\nu\left(\nabla\left(\boldsymbol{u}_{n+1}^h-\boldsymbol{u}_n^h\right),\nabla\boldsymbol{v}^h\right)+n_{\mathrm{skew}}\left(\boldsymbol{u}_{n+1},\boldsymbol{u}_{n+1},\boldsymbol{v}^h\right)\\
&\qquad -n_{\mathrm{skew}}\left(\boldsymbol{u}_{n+1}^h,\boldsymbol{u}_{n+1}^h,\boldsymbol{v}^h\right)-\left(\nabla\cdot\boldsymbol{v}^h,p-q^h\right)=0\quad\forall\ \boldsymbol{v}^h\in V_{\mathrm{div}}^h,q^h\in Q^h.
\end{aligned}
$$

Next, the decomposition (7.108) of the error is inserted and $\boldsymbol{v}^h = \boldsymbol{\phi}_{n+1}^h$ is chosen, leading to

$$\left(\partial_t \boldsymbol{u}_{n+1} - \frac{\boldsymbol{u}_{n+1}^h - \boldsymbol{u}_n^h}{\Delta t}, \boldsymbol{\phi}_{n+1}^h\right) + \nu\left(\nabla \boldsymbol{\eta}_{n+1}, \nabla \boldsymbol{\phi}_{n+1}^h\right) - \nu \left\|\nabla \boldsymbol{\phi}_{n+1}^h\right\|_{L^2(\Omega)}^2 \tag{7.109}$$
$$+ n_{\text{skew}}\left(\boldsymbol{u}_{n+1}, \boldsymbol{u}_{n+1}, \boldsymbol{\phi}_{n+1}^h\right) - n_{\text{skew}}\left(\boldsymbol{u}_{n+1}^h, \boldsymbol{u}_{n+1}^h, \boldsymbol{\phi}_{n+1}^h\right) - \left(\nabla \cdot \boldsymbol{\phi}_{n+1}^h, p - q^h\right) = 0$$

for all $q^h \in Q^h$. From the definition (4.54) of the Stokes projection and since $\boldsymbol{\phi}_{n+1}^h \in V_{\text{div}}^h$, it follows that $\left(\nabla \boldsymbol{\eta}_{n+1}, \nabla \boldsymbol{\phi}_{n+1}^h\right) = 0$. The left argument of the first term of (7.109) can be expanded in the form

$$\begin{aligned} &\partial_t \boldsymbol{u}_{n+1} - \frac{\boldsymbol{u}_{n+1}^h - \boldsymbol{u}_n^h}{\Delta t} \\ &= \partial_t \boldsymbol{u}_{n+1} - \partial_t I_{\text{St}}^h \boldsymbol{u}_{n+1} + \partial_t I_{\text{St}}^h \boldsymbol{u}_{n+1} - \frac{I_{\text{St}}^h \boldsymbol{u}_{n+1} - I_{\text{St}}^h \boldsymbol{u}_n}{\Delta t} \\ &\quad + \frac{I_{\text{St}}^h \boldsymbol{u}_{n+1} - I_{\text{St}}^h \boldsymbol{u}_n}{\Delta t} - \frac{\boldsymbol{u}_{n+1}^h - \boldsymbol{u}_n^h}{\Delta t} \\ &= \partial_t\left(\boldsymbol{u}_{n+1} - I_{\text{St}}^h \boldsymbol{u}_{n+1}\right) + \partial_t I_{\text{St}}^h \boldsymbol{u}_{n+1} - \frac{I_{\text{St}}^h \boldsymbol{u}_{n+1} - I_{\text{St}}^h \boldsymbol{u}_n}{\Delta t} - \frac{\boldsymbol{\phi}_{n+1}^h - \boldsymbol{\phi}_n^h}{\Delta t}. \end{aligned} \tag{7.110}$$

Inserting this decomposition in (7.109) gives a term of the form $\left(\boldsymbol{\phi}_{n+1}^h - \boldsymbol{\phi}_n^h, \boldsymbol{\phi}_{n+1}^h\right)$ which can be expanded in the same way as presented in (7.98). Then, one obtains from (7.109) the error equation

$$\begin{aligned} &\frac{1}{2\Delta t}\left(\left\|\boldsymbol{\phi}_{n+1}^h\right\|_{L^2(\Omega)}^2 + \left\|\boldsymbol{\phi}_{n+1}^h - \boldsymbol{\phi}_n^h\right\|_{L^2(\Omega)}^2 - \left\|\boldsymbol{\phi}_n^h\right\|_{L^2(\Omega)}^2\right) + \nu\left\|\nabla \boldsymbol{\phi}_{n+1}^h\right\|_{L^2(\Omega)}^2 \\ &= \left(\partial_t\left(\boldsymbol{u}_{n+1} - I_{\text{St}}^h \boldsymbol{u}_{n+1}\right), \boldsymbol{\phi}_{n+1}^h\right) + \left(\partial_t I_{\text{St}}^h \boldsymbol{u}_{n+1} - \frac{I_{\text{St}}^h \boldsymbol{u}_{n+1} - I_{\text{St}}^h \boldsymbol{u}_n}{\Delta t}, \boldsymbol{\phi}_{n+1}^h\right) \\ &\quad + n_{\text{skew}}\left(\boldsymbol{u}_{n+1}, \boldsymbol{u}_{n+1}, \boldsymbol{\phi}_{n+1}^h\right) - n_{\text{skew}}\left(\boldsymbol{u}_{n+1}^h, \boldsymbol{u}_{n+1}^h, \boldsymbol{\phi}_{n+1}^h\right) - \left(\nabla \cdot \boldsymbol{\phi}_{n+1}^h, p - q^h\right) \end{aligned} \tag{7.111}$$

for all $q^h \in Q^h$.

2. *Estimate all terms on the right hand-side of the error equation.* The first term on the right-hand side of (7.111) is bounded with the estimate of the dual pairing and Young's inequality (A.5)

$$\begin{aligned} \left(\partial_t\left(\boldsymbol{u}_{n+1} - I_{\text{St}}^h \boldsymbol{u}_{n+1}\right), \boldsymbol{\phi}_{n+1}^h\right) &\le \left\|\partial_t\left(\boldsymbol{u}_{n+1} - I_{\text{St}}^h \boldsymbol{u}_{n+1}\right)\right\|_{V'} \left\|\nabla \boldsymbol{\phi}_{n+1}^h\right\|_{L^2(\Omega)} \\ &\le \frac{4}{\nu}\left\|\partial_t\left(\boldsymbol{u}_{n+1} - I_{\text{St}}^h \boldsymbol{u}_{n+1}\right)\right\|_{V'}^2 + \frac{\nu}{16}\left\|\nabla \boldsymbol{\phi}_{n+1}^h\right\|_{L^2(\Omega)}^2. \end{aligned}$$

The estimate of the second term uses the Cauchy–Schwarz inequality (A.10), Poincaré's inequality (A.12), estimate (7.93), Young's inequality, and the commutation of temporal derivative and Stokes projection, see Remark 4.44,

$$\begin{aligned}
&\left(\partial_t I_{\mathrm{St}}^h \boldsymbol{u}_{n+1} - \frac{I_{\mathrm{St}}^h \boldsymbol{u}_{n+1} - I_{\mathrm{St}}^h \boldsymbol{u}_n}{\Delta t}, \boldsymbol{\phi}_{n+1}^h\right) \\
&\le \left\| \partial_t I_{\mathrm{St}}^h \boldsymbol{u}_{n+1} - \frac{I_{\mathrm{St}}^h \boldsymbol{u}_{n+1} - I_{\mathrm{St}}^h \boldsymbol{u}_n}{\Delta t} \right\|_{L^2(\Omega)} \left\| \boldsymbol{\phi}_{n+1}^h \right\|_{L^2(\Omega)} \\
&\le C(\Delta t)^{1/2} \left\| \partial_{tt} I_{\mathrm{St}}^h \boldsymbol{u} \right\|_{L^2\left(t_n,t_{n+1};L^2(\Omega)\right)} \left\| \nabla \boldsymbol{\phi}_{n+1}^h \right\|_{L^2(\Omega)} \\
&\le \frac{C\Delta t}{\nu} \left\| \partial_{tt} I_{\mathrm{St}}^h \boldsymbol{u} \right\|_{L^2\left(t_n,t_{n+1};L^2(\Omega)\right)}^2 + \frac{\nu}{16} \left\| \nabla \boldsymbol{\phi}_{n+1}^h \right\|_{L^2(\Omega)}^2 \\
&= \frac{C\Delta t}{\nu} \left\| I_{\mathrm{St}}^h \left(\partial_{tt} \boldsymbol{u}\right) \right\|_{L^2\left(t_n,t_{n+1};L^2(\Omega)\right)}^2 + \frac{\nu}{16} \left\| \nabla \boldsymbol{\phi}_{n+1}^h \right\|_{L^2(\Omega)}^2 .
\end{aligned}$$

The term with the pressure on the right-hand side of (7.111) is estimated in the same way as the corresponding term in the second step of the proof of Theorem 7.35

$$\left| \left(\nabla \cdot \boldsymbol{\phi}_{n+1}^h, p_{n+1} - q^h \right) \right| \le \frac{4}{\nu} \left\| p_{n+1} - q^h \right\|_{L^2(\Omega)}^2 + \frac{\nu}{16} \left\| \nabla \boldsymbol{\phi}_{n+1}^h \right\|_{L^2(\Omega)}^2 .$$

The nonlinear terms are decomposed in the form (6.65) and then the bounds of the arising three terms have the following form, compare the proof of Theorem 7.35,

$$\begin{aligned}
\left| n_{\mathrm{skew}} \left(\boldsymbol{\eta}_{n+1}, \boldsymbol{u}_{n+1}, \boldsymbol{\phi}_{n+1}^h \right) \right| &\le \frac{C}{\nu} \left\| \boldsymbol{\eta}_{n+1} \right\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{\eta}_{n+1} \right\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{u}_{n+1} \right\|_{L^2(\Omega)}^2 \\
&\quad + \frac{\nu}{16} \left\| \nabla \boldsymbol{\phi}_{n+1}^h \right\|_{L^2(\Omega)}^2 , \\
\left| n_{\mathrm{skew}} \left(\boldsymbol{\phi}_{n+1}^h, \boldsymbol{u}_{n+1}, \boldsymbol{\phi}_{n+1}^h \right) \right| &\le \frac{C}{\nu^3} \left\| \boldsymbol{\phi}_{n+1}^h \right\|_{L^2(\Omega)}^2 \left\| \nabla \boldsymbol{u}_{n+1} \right\|_{L^2(\Omega)}^4 + \frac{\nu}{16} \left\| \nabla \boldsymbol{\phi}_{n+1}^h \right\|_{L^2(\Omega)}^2 , \\
\left| n_{\mathrm{skew}} \left(\boldsymbol{u}_{n+1}^h, \boldsymbol{\eta}_{n+1}, \boldsymbol{\phi}_{n+1}^h \right) \right| &\le \frac{C}{\nu} \left\| \boldsymbol{u}_{n+1}^h \right\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{u}_{n+1}^h \right\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{\eta}_{n+1} \right\|_{L^2(\Omega)}^2 \\
&\quad + \frac{\nu}{16} \left\| \nabla \boldsymbol{\phi}_{n+1}^h \right\|_{L^2(\Omega)}^2 .
\end{aligned}$$

Now, all bounds are inserted in (7.111) and the second term on the left-hand side is estimated by zero from below. One gets

$$\begin{aligned}
&\frac{1}{2\Delta t} \left\| \boldsymbol{\phi}_{n+1}^h \right\|_{L^2(\Omega)}^2 + \frac{\nu}{2} \left\| \nabla \boldsymbol{\phi}_{n+1}^h \right\|_{L^2(\Omega)}^2 \\
&\le \frac{1}{2\Delta t} \left\| \boldsymbol{\phi}_{n}^h \right\|_{L^2(\Omega)}^2 + C \left[\frac{1}{\nu} \left(\left\| \partial_t \left(\boldsymbol{u}_{n+1} - I_{\mathrm{St}}^h \boldsymbol{u}_{n+1} \right) \right\|_{V'}^2 + \left\| p_{n+1} - q^h \right\|_{L^2(\Omega)}^2 \right.\right.
\end{aligned}$$

$$+\Delta t \left\| I_{\mathrm{St}}^h (\partial_{tt}\boldsymbol{u}) \right\|_{L^2(t_n,t_{n+1};L^2(\Omega))}^2 + \left\| \boldsymbol{\eta}_{n+1} \right\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{\eta}_{n+1} \right\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{u}_{n+1} \right\|_{L^2(\Omega)}^2$$

$$+ \left\| \boldsymbol{u}_{n+1}^h \right\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{u}_{n+1}^h \right\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{\eta}_{n+1} \right\|_{L^2(\Omega)}^2 \Bigg)$$

$$+ \frac{1}{\nu^3} \left\| \boldsymbol{\phi}_{n+1}^h \right\|_{L^2(\Omega)}^2 \left\| \nabla \boldsymbol{u}_{n+1} \right\|_{L^2(\Omega)}^4 \Bigg]$$

for all $q^h \in Q^h$. Taking the sum over all discrete times and using that the $L^2(\Omega)$ norms of the finite element remainders form a telescopic sum gives

$$\left\| \boldsymbol{\phi}_{N+1}^h \right\|_{L^2(\Omega)}^2 + \nu \Delta t \sum_{n=1}^{N+1} \left\| \nabla \boldsymbol{\phi}_n^h \right\|_{L^2(\Omega)}^2 \tag{7.112}$$

$$\leq \left\| \boldsymbol{\phi}_0^h \right\|_{L^2(\Omega)}^2 + C \Delta t \sum_{n=1}^{N+1} \mathcal{F}_n + \frac{C \Delta t}{\nu^3} \sum_{n=1}^{N+1} \left\| \boldsymbol{\phi}_n^h \right\|_{L^2(\Omega)}^2 \left\| \nabla \boldsymbol{u}_n \right\|_{L^2(\Omega)}^4 ,$$

with $\mathcal{F}_n$ representing all terms on the right-hand side that do not depend on $\boldsymbol{\phi}_n^h$.

3. *Application of a discrete Gronwall lemma.* Inequality (7.112) has exactly the form of (A.42) such that a discrete Gronwall inequality can be applied if all terms on the right-hand side of (7.112) are well defined. All norms of the solution and the Stokes projection for a given time t_n are well defined by the regularity assumptions on the solution. In addition, the stability estimate (4.55) of the Stokes projection yields

$$\sum_{n=1}^{N+1} \left\| I_{\mathrm{St}}^h (\partial_{tt}\boldsymbol{u}) \right\|_{L^2(t_{n-1},t_n;L^2(\Omega))}^2 = \left\| I_{\mathrm{St}}^h (\partial_{tt}\boldsymbol{u}) \right\|_{L^2(0,t_{N+1};L^2(\Omega))}^2$$

$$\leq C \left\| \partial_{tt}\boldsymbol{u} \right\|_{L^2(0,t_{N+1};H^1(\Omega))}^2 ,$$

which is finite by assumption (7.105). For the term with the finite element solution, one finds with the Cauchy–Schwarz inequality for sums (A.2) and the stability estimate (7.96)

$$\Delta t \sum_{n=1}^{N+1} \left\| \boldsymbol{u}_n^h \right\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{u}_n^h \right\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{\eta}_n \right\|_{L^2(\Omega)}^2$$

$$\leq \max_{n=1,\ldots,N} \left\| \boldsymbol{u}_n^h \right\|_{L^2(\Omega)} \Delta t \sum_{n=1}^{N+1} \left\| \nabla \boldsymbol{u}_n^h \right\|_{L^2(\Omega)} \left\| \nabla \boldsymbol{\eta}_n \right\|_{L^2(\Omega)}^2 \tag{7.113}$$

$$\leq \max_{n=1,\ldots,N} \left\|\boldsymbol{u}_n^h\right\|_{L^2(\Omega)} \left(\Delta t \sum_{n=1}^{N+1} \left\|\nabla \boldsymbol{u}_n^h\right\|_{L^2(\Omega)}^2\right)^{1/2} \left(\Delta t \sum_{n=1}^{N+1} \|\nabla \boldsymbol{\eta}_n\|_{L^2(\Omega)}^4\right)^{1/2}$$

$$\leq \frac{1}{\nu^{1/2}} \left(\left\|\boldsymbol{u}_0^h\right\|_{L^2(\Omega)}^2 + \frac{\Delta t}{\nu} \sum_{n=0}^{N} \left\|\boldsymbol{f}_{n+1}\right\|_{H^{-1}(\Omega)}^2\right) \left(\Delta t \sum_{n=1}^{N+1} \|\nabla \boldsymbol{\eta}_n\|_{L^2(\Omega)}^4\right)^{1/2}.$$

Thus, in the case that the time step is sufficiently small, i.e., if (7.106) is satisfied, the discrete Gronwall lemma, Lemma A.56, can be applied and one gets from (7.112), compare (A.43),

$$\left\|\boldsymbol{\phi}_{N+1}^h\right\|_{L^2(\Omega)}^2 + \nu \Delta t \sum_{n=1}^{N+1} \left\|\nabla \boldsymbol{\phi}_n^h\right\|_{L^2(\Omega)}^2$$

$$\leq \exp\left(\Delta t \sum_{n=1}^{N+1} \frac{\alpha_n}{1 - \Delta t \alpha_n}\right) \left(\left\|\boldsymbol{\phi}_0^h\right\|_{L^2(\Omega)}^2 + C \Delta t \sum_{n=1}^{N+1} \mathcal{F}_n'\right), \qquad (7.114)$$

with α_n defined by (7.106) and $\mathcal{F}_n'$ arises from $\mathcal{F}_n$ with inserting estimate (7.113).

4. *Application of the triangle inequality.* The application of the triangle inequality gives

$$\|\boldsymbol{e}_{N+1}\|_{L^2(\Omega)}^2 + \nu \Delta t \sum_{n=1}^{N+1} \|\nabla \boldsymbol{e}_n\|_{L^2(\Omega)}^2$$

$$\leq 2\left(\left\|\boldsymbol{\eta}_{N+1}\right\|_{L^2(\Omega)}^2 + \nu \Delta t \sum_{n=1}^{N+1} \|\nabla \boldsymbol{\eta}_n\|_{L^2(\Omega)}^2 + \left\|\boldsymbol{\phi}_{N+1}^h\right\|_{L^2(\Omega)}^2 + \nu \Delta t \sum_{n=1}^{N+1} \left\|\nabla \boldsymbol{\phi}_n^h\right\|_{L^2(\Omega)}^2\right).$$

Inserting (7.114) and the expression for $\mathcal{F}_n$ from (7.112) and (7.113) finishes the proof. ■

Remark 7.73 (On Estimate (7.107)*)*

- The error, the square root of the left-hand side of (7.107), is bounded by a first order term in time, namely the square root of the term $(\Delta t)^2 \left\|I_{\mathrm{St}}^h\left(\partial_{tt}\boldsymbol{u}\right)\right\|_{L^2\left(0,t_{N+1};L^2(\Omega)\right)}^2$. This first order convergence is the expected order for the backward Euler scheme.

- All other terms in the error bound contain interpolation errors in space. The most important ones are

$$\left(\Delta t \sum_{n=1}^{N+1} \left\|\nabla\left(\boldsymbol{u}_n - I_{\mathrm{St}}^h \boldsymbol{u}_n\right)\right\|_{L^2(\Omega)}^2\right)^{1/2}, \left(\Delta t \sum_{n=1}^{N+1} \left\|\nabla\left(\boldsymbol{u}_n - I_{\mathrm{St}}^h \boldsymbol{u}_n\right)\right\|_{L^2(\Omega)}^4\right)^{1/4},$$

$$\left(\inf_{q^h \in Q^h} \Delta t \sum_{n=1}^{N+1} \left\|p_n - q^h\right\|_{L^2(\Omega)}^2\right)^{1/2}.$$

  Error estimates for the Stokes projection are provided in (4.57). Note that the pressure term vanishes in this error bound since $p = 0$.
- The Stokes projection can be defined alternatively in the same way as described in Remark 7.38.
- Consider the situation $\Delta t \to 0$. Then the norms which are defined as sums over the discrete times become time-space norms, see Remark 7.71, e.g.,

$$\Delta t \sum_{n=1}^{N+1} \nu \left\|\nabla\left(\boldsymbol{u}_n - I_{\mathrm{St}}^h \boldsymbol{u}_n\right)\right\|_{L^2(\Omega)}^2 \to \nu \left\|\nabla\left(\boldsymbol{u} - I_{\mathrm{St}}^h \boldsymbol{u}\right)\right\|_{L^2\left(0,T;L^2(\Omega)\right)}^2.$$

  With the Poincaré inequality (A.12) and the Cauchy–Schwarz inequality for sums (A.2), one obtains

$$\begin{aligned}
&\Delta t \sum_{n=1}^{N+1} \left\|\boldsymbol{u}_n - I_{\mathrm{St}}^h \boldsymbol{u}_n\right\|_{L^2(\Omega)} \left\|\nabla\left(\boldsymbol{u}_n - I_{\mathrm{St}}^h \boldsymbol{u}_n\right)\right\|_{L^2(\Omega)} \left\|\nabla \boldsymbol{u}_n\right\|_{L^2(\Omega)}^2 \\
&\le \Delta t \sum_{n=1}^{N+1} \left\|\nabla\left(\boldsymbol{u}_n - I_{\mathrm{St}}^h \boldsymbol{u}_n\right)\right\|_{L^2(\Omega)}^2 \left\|\nabla \boldsymbol{u}_n\right\|_{L^2(\Omega)}^2 \\
&\le \left(\Delta t \sum_{n=1}^{N+1} \left\|\nabla\left(\boldsymbol{u}_n - I_{\mathrm{St}}^h \boldsymbol{u}_n\right)\right\|_{L^2(\Omega)}^4\right)^{1/2} \left(\Delta t \sum_{n=1}^{N+1} \left\|\nabla \boldsymbol{u}_n\right\|_{L^2(\Omega)}^4\right)^{1/2} \\
&\to \left\|\nabla\left(\boldsymbol{u} - I_{\mathrm{St}}^h \boldsymbol{u}\right)\right\|_{L^4\left(0,T;L^2(\Omega)\right)}^2 \left\|\nabla \boldsymbol{u}\right\|_{L^4\left(0,T;L^2(\Omega)\right)}^2.
\end{aligned}$$

  For the term with the exponential, it will be assumed that a bound of form (7.106) holds uniformly, i.e., there is a constant C_0 such that

$$C \frac{\left\|\nabla \boldsymbol{u}_n\right\|_{L^2(\Omega)}^4}{\nu^3} \Delta t \le C_0 < 1 \quad \forall\ t \in (0, T].$$

Then it is

$$\exp\left(\Delta t \sum_{n=1}^{N+1} \frac{\alpha_n}{1-\Delta t\alpha_n}\right) \le \exp\left(\frac{C}{\nu^3}\Delta t \sum_{n=1}^{N+1} \|\nabla \boldsymbol{u}_n\|_{L^2(\Omega)}^4\right) \tag{7.115}$$
$$\to \exp\left(\frac{C}{\nu^3}\|\nabla \boldsymbol{u}\|_{L^4(0,T;L^2(\Omega))}^4\right).$$

Finally, the term with $(\Delta t)^2 \left\| I_{\mathrm{St}}^h(\partial_{tt}\boldsymbol{u})\right\|_{L^2(0,t_{N+1};L^2(\Omega))}^2$ vanishes as $\Delta \to 0$.

Altogether, the terms of the error bound (7.107) converge to these of the bound (7.44) for the time-continuous case.

□

Remark 7.74 (Higher Regularity Assumptions) Similarly to the continuous-in-time case, different estimates of the convective term can be performed if a higher regularity of the solution $(\boldsymbol{u}, p)$ of (7.40) is assumed, compare Remark 7.39. In such cases, the time step restriction becomes weaker, e.g., using the assumptions of Remark 7.39 and estimate (7.58) leads to the requirement

$$\left(\frac{1}{2}\|\nabla \boldsymbol{u}_n\|_{L^\infty(\Omega)} + \frac{4}{\nu}\|\boldsymbol{u}_n\|_{L^\infty(\Omega)}^2\right)\Delta t < 1, \quad \forall\, n = 1,\ldots,N+1, \tag{7.116}$$

instead of (7.106). □

Remark 7.75 (Error Analysis for the Pressure) An error estimate for the pressure is based on the discrete inf-sup condition (3.51) and it starts in the same way, as, e.g., the estimate for the steady-state Navier–Stokes equations, compare the proof of Theorem 6.30. One obtains a formula of type (6.68) with the additional term

$$\left(\partial_t \boldsymbol{u}_{n+1} - \frac{\boldsymbol{u}_{n+1}^h - \boldsymbol{u}_n^h}{\Delta t}, \boldsymbol{v}^h\right)$$

in the numerator. The first component of this term can be decomposed in the form (7.110). Then, the hardest contribution to estimate is

$$\left(\frac{I_{\mathrm{St}}^h \boldsymbol{u}_{n+1} - \boldsymbol{u}_{n+1}^h}{\Delta t} - \frac{I_{\mathrm{St}}^h \boldsymbol{u}_n - \boldsymbol{u}_n^h}{\Delta t}, \boldsymbol{v}^h\right).$$

Estimating this term uses tools from Ayuso et al. (2005) that are also applied in the continuous-in-time case, compare the proof of Theorem 7.42. Details of this application can be found, e.g., in de Frutos et al. (2016b). Finally, it can be proved that

$$\Delta t \sum_{n=1}^{N+1} \left\|p_n - p_n^h\right\|_{L^2(\Omega)}^2$$

converges with the same order as the velocity errors bounded in (7.107). □

Remark 7.76 (Semi-implicit (IMEX) Euler Scheme) IMEX schemes approximate the equation at the discrete time t_{n+1} with a linear equation by using a convection field that can be computed from already known velocity fields, compare Remark 7.61. In the case of the Euler scheme, one simply takes the velocity field of the previous discrete time as convection field: Assume an equidistant time step, given $\boldsymbol{u}_n^h \in V^h$, compute $\left(\boldsymbol{u}_{n+1}^h, p_{n+1}^h\right) \in V^h \times Q^h$, $n = 0, 1, 2, \ldots$ by solving

$$\begin{aligned}&\frac{1}{\Delta t}\left(\boldsymbol{u}_{n+1}^h - \boldsymbol{u}_n^h, \boldsymbol{v}^h\right) + \nu\left(\nabla \boldsymbol{u}_{n+1}^h, \nabla \boldsymbol{v}^h\right) + n_{\text{skew}}\left(\boldsymbol{u}_n^h, \boldsymbol{u}_{n+1}^h, \boldsymbol{v}^h\right) \\ &-\left(\nabla\cdot\boldsymbol{v}^h, p_{n+1}^h\right) + \left(\nabla\cdot\boldsymbol{u}_{n+1}^h, q^h\right) = \langle \boldsymbol{f}_{n+1}, \boldsymbol{v}^h\rangle_{V',V}\ \forall\ \left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h,\end{aligned} \tag{7.117}$$

with $\boldsymbol{u}_0$ being an suitable approximation of the initial condition in the finite element velocity space.

With the scheme (7.117), one has to solve in each discrete time only one linear saddle point problem instead of a nonlinear saddle point problem. □

Lemma 7.77 (Existence, Uniqueness, and Stability of the Finite Element Solution) *Let $V^h \times Q^h$ be conforming finite element spaces which satisfy the discrete inf-sup condition* (3.51)*. If $\boldsymbol{u}_n^h \in V^h$, then* (7.117) *has a unique solution and for all $N > 0$, it holds the stability estimate*

$$\left\|\boldsymbol{u}_{N+1}^h\right\|_{L^2(\Omega)}^2 + \nu\Delta t\sum_{n=0}^N \left\|\nabla \boldsymbol{u}_{n+1}^h\right\|_{L^2(\Omega)}^2 \le \left\|\boldsymbol{u}_0^h\right\|_{L^2(\Omega)}^2 + \frac{\Delta t}{\nu}\sum_{n=0}^N \left\|\boldsymbol{f}_{n+1}\right\|_{H^{-1}(\Omega)}^2. \tag{7.118}$$

Proof The existence and uniqueness of the velocity solution is proved with Theorem of Lax–Milgram, see Theorem B.4, which can be applied since (7.117) is a linear problem. To this end, consider all terms with $\boldsymbol{u}_{n+1}^h$ on the left-hand side of (7.117) in V_{div}^h and denote the corresponding bilinear form with $a(\cdot,\cdot)$. Then both terms with the divergence vanish. The coercivity follows with a straightforward calculation, using that the convective term vanishes if the second and third argument are the same, see (6.26),

$$a\left(\boldsymbol{v}^h, \boldsymbol{v}^h\right) = \frac{1}{\Delta t}\left\|\boldsymbol{v}^h\right\|_{L^2(\Omega)}^2 + \nu\left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)}^2 \ge \nu\left\|\nabla \boldsymbol{v}^h\right\|_{L^2(\Omega)}^2 = \nu\left\|\nabla \boldsymbol{v}^h\right\|_V^2.$$

The bilinear form is bounded, since one obtains with the Cauchy–Schwarz inequality (A.10), estimate (6.30), the Poincaré inequality (A.12), and $\left\|\nabla \boldsymbol{u}_n^h\right\|_{L^2(\Omega)} < \infty$, because by assumption $\boldsymbol{u}_n^h \in V^h \subset V$,

$$
\begin{aligned}
a\left(\boldsymbol{v}^h, \boldsymbol{w}^h\right) &\le \frac{1}{\Delta t}\left\|\boldsymbol{v}^h\right\|_{L^2(\Omega)}\left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)} + \nu\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{w}^h\right\|_{L^2(\Omega)} \\
&\quad + \left\|\nabla\boldsymbol{u}_n^h\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{w}^h\right\|_{L^2(\Omega)} \\
&\le C\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{w}^h\right\|_{L^2(\Omega)}.
\end{aligned}
$$

The boundedness of the right-hand side of (7.117) follows by the estimate of the dual pairing, the Cauchy–Schwarz inequality, and the Poincaré inequality

$$
\begin{aligned}
&\langle \boldsymbol{f}_{n+1}, \boldsymbol{v}^h\rangle_{V',V} + \frac{1}{\Delta t}\left(\boldsymbol{u}_n^h, \boldsymbol{v}^h\right) \\
&\le \left\|\boldsymbol{f}_{n+1}\right\|_{H^{-1}(\Omega)}\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)} + \frac{1}{\Delta t}\left\|\boldsymbol{u}_n^h\right\|_{L^2(\Omega)}\left\|\boldsymbol{v}^h\right\|_{L^2(\Omega)} \\
&\le \left(\left\|\boldsymbol{f}_{n+1}\right\|_{H^{-1}(\Omega)} + \frac{C}{\Delta t}\left\|\nabla\boldsymbol{u}_n^h\right\|_{L^2(\Omega)}\right)\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)} \le C\left\|\nabla\boldsymbol{v}^h\right\|_{L^2(\Omega)}.
\end{aligned}
$$

Thus, the assumptions of the Theorem of Lax–Milgram are satisfied, such that the existence and uniqueness of a finite element velocity solution follows. The existence and uniqueness of a finite element pressure solution is a direct consequence of the fact that the finite element spaces satisfy the discrete inf-sup condition.

The proof of the stability is exactly the same as for the fully implicit Euler scheme, see the proof of Lemma 7.69. ■

Theorem 7.78 (Finite Element Error Estimate for the Velocity Computed with the IMEX Euler Scheme in Combination with the Galerkin Finite Element Method) *Let the same assumptions on the data of the problem hold as stated in Theorem 7.72. Assume the following regularities for the solution* $(\boldsymbol{u}, p)$ *of the Navier–Stokes equations* (7.40)

$$
\begin{aligned}
&\partial_{tt}\boldsymbol{u} \in L^4\left(0,T;L^2(\Omega)\right), \qquad \partial_t\boldsymbol{u} \in L^4\left(0,T;L^2(\Omega)\right), \\
&\boldsymbol{u} \in L^4\left(0,T;W^{1,\infty}(\Omega)\right), \quad p \in L^2\left(0,T;L^2(\Omega)\right).
\end{aligned} \tag{7.119}
$$

Consider the IMEX Euler scheme (7.117) *with equidistant time steps and with pairs of finite element spaces that satisfy the discrete inf-sup condition* (3.51)*. Then the following error estimate is valid for all* $N > 0$

$$
\begin{aligned}
&\left\|\boldsymbol{u}_{N+1} - \boldsymbol{u}_{N+1}^h\right\|_{L^2(\Omega)}^2 + \nu\Delta t\sum_{n=1}^{N+1}\left\|\nabla\left(\boldsymbol{u}_n - \boldsymbol{u}_n^h\right)\right\|_{L^2(\Omega)}^2 \\
&\le C\Bigg\{\left\|\boldsymbol{u}_{N+1} - I_{\mathrm{St}}^h\boldsymbol{u}_{N+1}\right\|_{L^2(\Omega)}^2 + \nu\Delta t\sum_{n=1}^{N+1}\left\|\nabla\left(\boldsymbol{u}_n - I_{\mathrm{St}}^h\boldsymbol{u}_n\right)\right\|_{L^2(\Omega)}^2
\end{aligned}
$$

$$+ \exp\left(C\frac{\Delta t}{\nu}\sum_{n=0}^{N}\|\boldsymbol{u}_{n+1}\|^2_{W^{1,\infty}(\Omega)}\right)\left(\left\|\boldsymbol{u}_0^h - I_{\mathrm{St}}^h\boldsymbol{u}(0)\right\|^2_{L^2(\Omega)} + \frac{\Delta t}{\nu}\sum_{n=1}^{N+1}\mathcal{F}_n\right)\Bigg\}, \tag{7.120}$$

with $\mathcal{F}_n$ defined to be the terms in parentheses of (7.121) below, which contains approximation errors in time and space. In (7.120), $I_{\mathrm{St}}^h\boldsymbol{u}_n$ is the Stokes projection of $\boldsymbol{u}_n$ with $p = 0$, see (4.54), for which $\partial_t I_{\mathrm{St}}^h\boldsymbol{u} \in L^2(0, T; V')$ is assumed.

Proof The proof of this theorem follows the proof of Theorem 7.72.

In the first step, only the difference of the nonlinear convective terms in the error equation looks different. To estimate this difference in the second step, it is rewritten in the following way

$$\begin{aligned}
&n_{\mathrm{skew}}\left(\boldsymbol{u}_{n+1}, \boldsymbol{u}_{n+1}, \boldsymbol{\phi}_{n+1}^h\right) - n_{\mathrm{skew}}\left(\boldsymbol{u}_n^h, \boldsymbol{u}_{n+1}^h, \boldsymbol{\phi}_{n+1}^h\right)\\
&= n_{\mathrm{skew}}\left(\boldsymbol{u}_{n+1}, \boldsymbol{u}_{n+1}, \boldsymbol{\phi}_{n+1}^h\right) - n_{\mathrm{skew}}\left(\boldsymbol{u}_n, \boldsymbol{u}_{n+1}, \boldsymbol{\phi}_{n+1}^h\right) + n_{\mathrm{skew}}\left(\boldsymbol{u}_n, \boldsymbol{u}_{n+1}, \boldsymbol{\phi}_{n+1}^h\right)\\
&\quad - n_{\mathrm{skew}}\left(\boldsymbol{u}_n^h, \boldsymbol{u}_{n+1}, \boldsymbol{\phi}_{n+1}^h\right) + n_{\mathrm{skew}}\left(\boldsymbol{u}_n^h, \boldsymbol{u}_{n+1}, \boldsymbol{\phi}_{n+1}^h\right) - n_{\mathrm{skew}}\left(\boldsymbol{u}_n^h, \boldsymbol{u}_{n+1}^h, \boldsymbol{\phi}_{n+1}^h\right)\\
&= n_{\mathrm{skew}}\left(\boldsymbol{u}_{n+1} - \boldsymbol{u}_n, \boldsymbol{u}_{n+1}, \boldsymbol{\phi}_{n+1}^h\right) + n_{\mathrm{skew}}\left(\boldsymbol{\eta}_n, \boldsymbol{u}_{n+1}, \boldsymbol{\phi}_{n+1}^h\right) - n_{\mathrm{skew}}\left(\boldsymbol{\phi}_n^h, \boldsymbol{u}_{n+1}, \boldsymbol{\phi}_{n+1}^h\right)\\
&\quad + n_{\mathrm{skew}}\left(\boldsymbol{u}_n^h, \boldsymbol{\eta}_{n+1}, \boldsymbol{\phi}_{n+1}^h\right) - n_{\mathrm{skew}}\left(\boldsymbol{u}_n^h, \boldsymbol{\phi}_{n+1}^h, \boldsymbol{\phi}_{n+1}^h\right).
\end{aligned}$$

The last term vanishes because of the skew-symmetry (6.26). The second and fourth term are estimated in a standard way, applying (6.41) and Young's inequality (A.5),

$$\begin{aligned}
\left|n_{\mathrm{skew}}\left(\boldsymbol{\eta}_n, \boldsymbol{u}_{n+1}, \boldsymbol{\phi}_{n+1}^h\right)\right| &\le \frac{C}{\nu}\|\boldsymbol{\eta}_n\|_{L^2(\Omega)}\|\nabla\boldsymbol{\eta}_n\|_{L^2(\Omega)}\|\nabla\boldsymbol{u}_{n+1}\|^2_{L^2(\Omega)}\\
&\quad + \frac{\nu}{16}\left\|\nabla\boldsymbol{\phi}_{n+1}^h\right\|^2_{L^2(\Omega)},\\
\left|n_{\mathrm{skew}}\left(\boldsymbol{u}_n^h, \boldsymbol{\eta}_{n+1}, \boldsymbol{\phi}_{n+1}^h\right)\right| &\le \frac{C}{\nu}\left\|\boldsymbol{u}_n^h\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{u}_n^h\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{\eta}_{n+1}\right\|^2_{L^2(\Omega)}\\
&\quad + \frac{\nu}{16}\left\|\nabla\boldsymbol{\phi}_{n+1}^h\right\|^2_{L^2(\Omega)}.
\end{aligned}$$

For the third term, estimate (6.40) is applied, using the assumed regularity of the velocity and Poincaré's inequality (A.12),

$$\begin{aligned}
\left|n_{\mathrm{skew}}\left(\boldsymbol{\phi}_n^h, \boldsymbol{u}_{n+1}, \boldsymbol{\phi}_{n+1}^h\right)\right| &\le C\left\|\boldsymbol{\phi}_n^h\right\|_{L^2(\Omega)}\|\boldsymbol{u}_{n+1}\|_{W^{1,\infty}(\Omega)}\left\|\nabla\boldsymbol{\phi}_{n+1}^h\right\|_{L^2(\Omega)}\\
&\le \frac{C}{\nu}\left\|\boldsymbol{\phi}_n^h\right\|^2_{L^2(\Omega)}\|\boldsymbol{u}_{n+1}\|^2_{W^{1,\infty}(\Omega)} + \frac{\nu}{16}\left\|\nabla\boldsymbol{\phi}_{n+1}^h\right\|^2_{L^2(\Omega)}.
\end{aligned}$$

The estimate of the first term starts in the same way

$$n_{\text{skew}}\left(\boldsymbol{u}_{n+1}-\boldsymbol{u}_n,\boldsymbol{u}_{n+1},\boldsymbol{\phi}_{n+1}^h\right) \\ \le \frac{C}{\nu}\left\|\boldsymbol{u}_{n+1}-\boldsymbol{u}_n\right\|_{L^2(\Omega)}^2\left\|\boldsymbol{u}_{n+1}\right\|_{W^{1,\infty}(\Omega)}^2+\frac{\nu}{16}\left\|\nabla\boldsymbol{\phi}_{n+1}^h\right\|_{L^2(\Omega)}^2.$$

Then, the triangle inequality and the consistency error estimate (7.93) are applied to get

$$\begin{aligned}\left\|\boldsymbol{u}_{n+1}-\boldsymbol{u}_n\right\|_{L^2(\Omega)}^2 &\le 2\left\|\Delta t\partial_t\boldsymbol{u}_{n+1}-\boldsymbol{u}_{n+1}-\boldsymbol{u}_n\right\|_{L^2(\Omega)}^2+2\left(\Delta t\right)^2\left\|\partial_t\boldsymbol{u}_{n+1}\right\|_{L^2(\Omega)}^2 \\ &\le 2\left(\Delta t\right)^3\left\|\partial_{tt}\boldsymbol{u}\right\|_{L^2\left(t_n,t_{n+1};L^2(\Omega)\right)}^2+2\left(\Delta t\right)^2\left\|\partial_t\boldsymbol{u}_{n+1}\right\|_{L^2(\Omega)}^2.\end{aligned}$$

Inserting all estimates in the error equation and summing over all discrete times gives

$$\begin{aligned}&\left\|\boldsymbol{\phi}_{N+1}^h\right\|_{L^2(\Omega)}^2+\nu\Delta t\sum_{n=1}^{N+1}\left\|\nabla\boldsymbol{\phi}_n^h\right\|_{L^2(\Omega)}^2 \\ &\le\left\|\boldsymbol{\phi}_0^h\right\|_{L^2(\Omega)}^2+\frac{C\Delta t}{\nu}\sum_{n=1}^{N+1}\Bigg(\left\|\partial_t\left(\boldsymbol{u}_n-I_{\text{St}}^h\boldsymbol{u}_n\right)\right\|_{V'}^2 \\ &\quad+\left\|\boldsymbol{\eta}_{n-1}\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{\eta}_{n-1}\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{u}_n\right\|_{L^2(\Omega)}^2 \\ &\quad+\left\|\boldsymbol{u}_{n-1}^h\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{u}_{n-1}^h\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{\eta}_n\right\|_{L^2(\Omega)}^2+(\Delta t)^2\left\|\partial_t\boldsymbol{u}_n\right\|_{L^2(\Omega)}^2\left\|\boldsymbol{u}_n\right\|_{W^{1,\infty}(\Omega)}^2 \\ &\quad+(\Delta t)^3\left\|\partial_{tt}\boldsymbol{u}\right\|_{L^2\left(t_{n-1},t_n;L^2(\Omega)\right)}^2\left\|\boldsymbol{u}_n\right\|_{W^{1,\infty}(\Omega)}^2+\Delta t\left\|I_{\text{St}}^h\left(\partial_{tt}\boldsymbol{u}\right)\right\|_{L^2\left(t_{n-1},t_n;L^2(\Omega)\right)}^2 \\ &\quad+\left\|p_n-q^h\right\|_{L^2(\Omega)}^2\Bigg)+\frac{C\Delta t}{\nu}\sum_{n=0}^{N}\left\|\boldsymbol{\phi}_n^h\right\|_{L^2(\Omega)}^2\left\|\boldsymbol{u}_{n+1}\right\|_{W^{1,\infty}(\Omega)}^2.\end{aligned} \tag{7.121}$$

This inequality is of the same type as inequality (A.44). Hence, the version (A.45) of the discrete Gronwall lemma can be applied if all terms on the right-hand side of (7.121) are finite. This property follows from the assumptions on the solution (7.119), e.g., one gets with the Cauchy–Schwarz inequality for sums (A.2)

$$\begin{aligned}&\Delta t\sum_{n=1}^{N+1}\left\|\partial_t\boldsymbol{u}_n\right\|_{L^2(\Omega)}^2\left\|\boldsymbol{u}_n\right\|_{W^{1,\infty}(\Omega)}^2 \\ &\le\left(\sum_{n=1}^{N+1}\Delta t\left\|\partial_t\boldsymbol{u}_n\right\|_{L^2(\Omega)}^4\right)^{1/2}\left(\sum_{n=1}^{N+1}\Delta t\left\|\boldsymbol{u}_n\right\|_{W^{1,\infty}(\Omega)}^4\right)^{1/2}.\end{aligned} \tag{7.122}$$

From Remark 7.71, it follows that the first term on the right-hand side is an approximation of $\|\partial_t \boldsymbol{u}\|^2_{L^4(0,t_{N+1};L^2(\Omega))}$ and the second term approximates $\|\boldsymbol{u}\|^2_{L^4(0,t_{N+1};W^{1,\infty}(\Omega))}$. In the same way, one gets

$$
\begin{aligned}
&(\Delta t)^{1/2} \sum_{n=1}^{N+1} \|\partial_{tt}\boldsymbol{u}\|^2_{L^2(t_{n-1},t_n;L^2(\Omega))} \|\boldsymbol{u}_n\|^2_{W^{1,\infty}(\Omega)} \\
&\quad \le \|\partial_{tt}\boldsymbol{u}\|^4_{L^2(0,t_{N+1};L^2(\Omega))} \left(\sum_{n=1}^{N+1} \Delta t \, \|\boldsymbol{u}_n\|^4_{W^{1,\infty}(\Omega)} \right)^{1/2}. \qquad (7.123)
\end{aligned}
$$

For the other terms on the right-hand side of (7.121), the same reasoning applies as in the proof of Theorem 7.72. The application of the discrete Gronwall inequality (A.45), which does not require a restriction on the length of the time step, gives

$$
\begin{aligned}
&\left\|\boldsymbol{\phi}^h_{N+1}\right\|^2_{L^2(\Omega)} + \nu \Delta t \sum_{n=1}^{N+1} \left\|\nabla \boldsymbol{\phi}^h_n\right\|^2_{L^2(\Omega)} \\
&\quad \le \exp\left(\frac{C\Delta t}{\nu} \sum_{n=0}^{N} \|\boldsymbol{u}_{n+1}\|^2_{W^{1,\infty}(\Omega)} \right) \left(\left\|\boldsymbol{\phi}^h_0\right\|^2_{L^2(\Omega)} + \frac{C\Delta t}{\nu} \sum_{n=1}^{N+1} \mathcal{F}_n \right),
\end{aligned}
$$

where $\mathcal{F}_n$ contains the terms in the parentheses of (7.121). The final step of the proof is the application of the triangle inequality. ■

Remark 7.79 (On Estimate (7.120))

- The approximation errors in the error bound (7.120) are of the same order as in the bound (7.107) for the fully implicit Euler scheme. Thus, asymptotically the fully implicit scheme and the IMEX scheme are of the same order in time and space. A concrete example that studies the accuracy and the efficiency of IMEX schemes is presented in Example 7.62.
- Inspecting estimates (7.122) and (7.123), it can be seen that increasing the assumptions on the temporal regularity of $\partial_t \boldsymbol{u}$ and $\partial_{tt}\boldsymbol{u}$, e.g., to $\partial_t \boldsymbol{u}, \partial_{tt}\boldsymbol{u} \in L^\infty\left(0,T;L^2(\Omega)\right)$, allows to reduce the regularity assumptions with respect to time on $\boldsymbol{u}$ to $\boldsymbol{u} \in L^2\left(0,T;W^{1,\infty}(\Omega)\right)$.
- In order to facilitate the comparison of the error bounds for the fully implicit scheme and the IMEX scheme, the exponential term in the bound for the implicit scheme should be estimated as in (7.115). For a fair comparison, similar regularity assumptions on the solution should be considered, e.g., the assumptions used in Remark 7.74 and (7.119).

  The bound (7.120) of the IMEX scheme is better in the respect that there is no restriction on the length of the time step. This property comes from the use of the discrete Gronwall inequality (A.45) for the IMEX scheme instead of (A.43).

However, a severe time step restriction for the fully implicit scheme is usually not observed in practice.

In the expression (7.116) that enters the estimate of the exponential for the fully implicit scheme, the term with the highest assumed regularity, $\|\nabla \boldsymbol{u}_n\|_{L^\infty(\Omega)}$, appears only linearly whereas this term appears quadratically in (7.120) and it is scaled with ν^{-1}. In practice, this norm is expected to be large such that in this respect the error bound for the fully implicit scheme is better.

□

Remark 7.80 (Analysis of Temporal Discretizations Applied to the Navier–Stokes Equations)

- Finite element error analysis for the fully discrete Navier–Stokes equations with Crank–Nicolson scheme can be found in Heywood and Rannacher (1990) and Bause (1997). Among other results, second order convergence in time of the velocity error $\left\|\boldsymbol{u}_n - \boldsymbol{u}_n^h\right\|_{L^2(\Omega)}$ for all time steps t_n was proved.
- The fractional-step θ-scheme was investigated analytically in Klouček and Rys (1994) and Müller-Urbaniak (1993). It was shown in Klouček and Rys (1994) that the solution computed with this scheme converges to the solution of the Navier–Stokes equation, whose uniqueness was assumed, as the mesh width and the time step length tend to zero. A second order error estimate with respect to time similar to the Crank–Nicolson scheme was proved in Müller-Urbaniak (1993).
- An IMEX Crank–Nicolson scheme for the Navier–Stokes equations was already analyzed in Baker (1976). A more refined finite element error analysis can be found in Ingram (2013b). For a so-called stabilized IMEX Crank–Nicolson scheme, CNLE(stab), see (7.90), stability estimates were derived in Ingram (2013a).
- An error analysis for a IMEX versions of the BDF2 scheme applied to the Navier–Stokes equations can be found already in Girault and Raviart (1979, Chap. V, Sect. 3.3). The analysis was performed for the space-continuous setting. Optimal second order estimates for

$$\sup_{0\le n\le N} \left\|\left(\boldsymbol{u}-\boldsymbol{u}^h\right)(t_n)\right\|_{L^2(\Omega)} \quad \text{and} \quad \Delta t^{1/2}\sum_{n=0}^{N-2}\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{u}^h\right)(t_n)\right\|_{L^2(\Omega)} \tag{7.124}$$

were proved with the assumption of sufficiently high regularity of the solution.

Under more realistic regularity assumptions, the BDF2 scheme was analyzed in Emmrich (2004a,b). In Emmrich (2004b), the existence and stability of a solution was studied as well as the convergence to a weak solution of the Navier–Stokes equations. A finite element error analysis, also in the space-continuous setting, was performed in Emmrich (2004a), leading to second order convergence for the time-weighted velocity in the norms (7.124) (with the velocity replaced by the time-weighted velocity) and first order convergence of the time-weighted

pressure in the norm that corresponds to the first norm of (7.124), both in the fully implicit case. For an IMEX approach, using an extrapolation of the previous two solutions, only reduced orders of convergence were proved. The analysis of the results for the fully implicit scheme requires sufficiently small time steps, depending strongly on ν^{-1}, whereas the convergence of the IMEX scheme could be proved without time step restriction.
- A numerical analysis for the fully discrete schemes with a post-processing step that improves the order of convergence in space can be found in de Frutos et al. (2008). This paper presents the analysis for the backward Euler scheme and for BDF2.

□

Remark 7.81 (Discretizations with Stabilizations) There are several papers with finite element error analysis of stabilized discretizations of time-dependent incompressible flow problems. Both, stabilizations with respect to the violation of the discrete inf-sup condition (3.51) and with respect to dominating convection were studied.

- The PSPG method, see Sect. 4.5.1, was analyzed for the evolutionary Stokes equations

$$\begin{aligned} \partial_t \boldsymbol{u} - \nu\Delta\boldsymbol{u} + \nabla p &= \boldsymbol{f} \quad \text{in } (0,T]\times\Omega, \\ \nabla\cdot\boldsymbol{u} &= 0 \quad \text{in } (0,T]\times\Omega, \\ \boldsymbol{u}(0,\cdot) &= \boldsymbol{u}_0 \text{ in } \Omega, \end{aligned} \tag{7.125}$$

in Burman and Fernández (2011) and John and Novo (2015). In Burman and Fernández (2011), the fully discrete scheme was studied and optimal convergence was proved, in the case of higher order finite element functions for sufficiently large time steps. Optimal error estimates without time step restriction were derived in John and Novo (2015), such that both the fully discrete case and the continuous-in-time case are covered.
- The analysis of the transient Stokes equations (7.125) presented in Burman and Fernández (2008) covers other stabilizations with respect to the discrete inf-sup condition, like the method from Brezzi and Pitkäranta (1984), Remark 4.110, the method proposed in Codina and Blasco (1997), Remark 4.111, the method developed in Dohrmann and Bochev (2004), Remark 4.113, the CIP method, Remark 5.49, and the LPS method, Remark 5.53.
- A numerical analysis of the grad-div stabilization for the evolutionary Oseen equations, i.e., (5.1) with the additional term $\partial_t \boldsymbol{u}$, is presented in de Frutos et al. (2016b). In both the continuous-in-time and the fully discrete case (with backward Euler scheme, Crank–Nicolson scheme, and BDF2), error bounds were derived that do not depend explicitly on the inverse of the viscosity. The analysis assumed a sufficiently smooth solution. The results were extended to the instationary Navier–Stokes equations in de Frutos et al. (2016a), covering

also the situation that the nonlocal compatibility conditions at the initial time, compare Remark 7.30, are not satisfied.

- The CIP method, Remark 5.49, applied with equal-order pairs of finite element spaces was analyzed in Burman and Fernández (2007) for the continuous-in-time case. Assuming a sufficiently smooth solution of the incompressible Navier–Stokes equations (7.1), error bounds were derived that do not depend on inverse powers of the viscosity.
- A finite element error analysis of the LPS method, see Sect. 5.4 starting with Remark 5.52, applied to the time-dependent Navier–Stokes equations was presented in Arndt et al. (2015). Inf-sup stable pairs of finite element spaces and the continuous-in-time case were considered. In the case of a sufficiently smooth solution, error bounds were proved where the Gronwall constant does not depend on the viscosity.

□

7.5 Approaches Decoupling Velocity and Pressure: Projection Methods

Remark 7.82 (Motivation and General Ideas) The motivation for the construction of projection methods was the wish to obtain simple schemes for simulating the time-dependent Navier–Stokes equations (7.1) that require only the solution of scalar linear systems of equations for each component of the velocity and the pressure but not of saddle point problems.

To this end, the Navier–Stokes equations (7.1) are decoupled such that separate equations for velocity and pressure are obtained. Let an approximation of the time derivative by a q-step scheme be given, for simplicity of presentation with an equidistant time step,

$$\partial_t \boldsymbol{u}\,(t_{n+1}) \approx \frac{1}{\Delta t}\left(\tau_q \boldsymbol{u}_{n+1} + \sum_{j=0}^{q-1} \tau_{q-1-j}\boldsymbol{u}_{n-j}\right), \quad \sum_{j=0}^{q} \tau_j = 0. \tag{7.126}$$

The equation for an intermediate velocity comes from the momentum balance of the Navier–Stokes equations. Given $\hat{p} \in L^2(\Omega)$ or $\nabla \hat{p}$ and the initial condition $\tilde{\boldsymbol{u}}(0) = \boldsymbol{u}(0) = \boldsymbol{u}_0$, it has the form

$$\begin{aligned} \frac{1}{\Delta t}\left(\tau_q \tilde{\boldsymbol{u}}_{n+1} + \sum_{j=0}^{q-1} \tau_{q-1-j}\boldsymbol{u}_{n-j}\right) & \\ -\nu \Delta \tilde{\boldsymbol{u}}_{n+1} + (\boldsymbol{u}_n \cdot \nabla)\tilde{\boldsymbol{u}}_{n+1} &= \boldsymbol{f}_{n+1} - \nabla \hat{p} \text{ in } \Omega, \\ \tilde{\boldsymbol{u}}_{n+1} &= \boldsymbol{0} \qquad\quad \text{on } \Gamma. \end{aligned} \tag{7.127}$$

To obtain only linear systems of equations, the convective term in (7.127) has to be treated semi-implicitly or even explicitly. Instead of $(\boldsymbol{u}_n \cdot \nabla)\tilde{\boldsymbol{u}}_{n+1}$, also the semi-implicit and explicit forms

$$(\tilde{\boldsymbol{u}}_n \cdot \nabla)\tilde{\boldsymbol{u}}_{n+1} \quad \text{or} \quad (\tilde{\boldsymbol{u}}_n \cdot \nabla)\tilde{\boldsymbol{u}}_n \quad \text{or} \quad (\boldsymbol{u}_n \cdot \nabla)\boldsymbol{u}_n$$

or even extrapolations from former time steps like

$$2(\tilde{\boldsymbol{u}}_n \cdot \nabla)\tilde{\boldsymbol{u}}_n - (\tilde{\boldsymbol{u}}_{n-1} \cdot \nabla)\tilde{\boldsymbol{u}}_{n-1} \tag{7.128}$$

can be used in all schemes presented in this section, see also Remark 7.100.

The intermediate velocity field $\tilde{\boldsymbol{u}}$ will be in general not divergence-free. For this reason, a correction step

$$\begin{aligned} \frac{1}{\Delta t}\left(\tau_q \boldsymbol{u}_{n+1} - \tau_q \tilde{\boldsymbol{u}}_{n+1}\right) + \nabla\varphi\left(\tilde{\boldsymbol{u}}_{n+1}\right) + \nabla p = \nabla\hat{p} &\text{ in } \Omega, \\ \nabla \cdot \boldsymbol{u}_{n+1} = 0 \quad &\text{in } \Omega, \\ \boldsymbol{u}_{n+1} \cdot \boldsymbol{n} = 0 \quad &\text{on } \Gamma, \end{aligned} \tag{7.129}$$

will be performed, where $\varphi(\cdot)$ is a given function with values in $L^2(\Omega)$. In Lemma 7.83, it will become clear that (7.129) describes the $L^2(\Omega)$ projection of $\tilde{\boldsymbol{u}}_{n+1}$ onto a velocity $\boldsymbol{u}_{n+1}$ in $H_{\text{div}}(\Omega)$. The examples given below will illustrate that (7.129) can be transformed into an equation for the pressure.

Note that (7.129) is only a first order partial differential equation with respect to space for $\boldsymbol{u}_{n+1}$. Hence, to define a well-posed problem, it is not possible to use the same boundary condition for $\boldsymbol{u}_{n+1}$ as given for the velocity in the Navier–Stokes equations, which are of second order with respect to space.

Thus, the definition of projection methods involves two velocity fields: $\tilde{\boldsymbol{u}}_{n+1}$ satisfies the boundary conditions but it is not divergence-free whereas $\boldsymbol{u}_{n+1}$ is divergence-free but it does not satisfy the boundary conditions given in the Navier–Stokes equations.

Adding (7.127) and (7.129) gives

$$\begin{aligned} \frac{1}{\Delta t}\left(\tau_q \tilde{\boldsymbol{u}}_{n+1} + \sum_{j=0}^{q-1} \tau_{q-1-j}\boldsymbol{u}_{n-j}\right) \\ -\nu\Delta\tilde{\boldsymbol{u}}_{n+1} + (\boldsymbol{u}_n \cdot \nabla)\tilde{\boldsymbol{u}}_{n+1} + \nabla\varphi\left(\tilde{\boldsymbol{u}}_{n+1}\right) + \nabla p = \boldsymbol{f}_{n+1} &\text{ in } \Omega, \\ \tilde{\boldsymbol{u}}_{n+1} = \boldsymbol{0} \quad &\text{on } \Gamma. \end{aligned}$$

One can see that an error is committed in applying the splitting, whose size depends in particular on $\nabla\varphi\left(\tilde{\boldsymbol{u}}_{n+1}\right)$.

A survey of projection methods for incompressible flow problems can be found in Guermond et al. (2006). Although projection schemes aim to facilitate the numerical solution of the incompressible Navier–Stokes equations, their analysis turns out to be more involved than the analysis for coupled schemes. Here, only the most important results concerning their convergence will be mentioned. For their proofs, it is referred to the literature. □

Lemma 7.83 (Projection Property of (7.129)) *The velocity* $\boldsymbol{u}_{n+1}$ *computed with (7.129) is the* $L^2(\Omega)$ *projection of* $\tilde{\boldsymbol{u}}_{n+1}$ *into* $H_{\text{div}}(\Omega)$*, i.e.,* $\boldsymbol{u}_{n+1}$ *is the Helmholtz projection of* $\tilde{\boldsymbol{u}}_{n+1}$*. It is* $\|\boldsymbol{u}_{n+1}\|_{L^2(\Omega)} \le \|\tilde{\boldsymbol{u}}_{n+1}\|_{L^2(\Omega)}$.

Proof Let $\boldsymbol{v} \in H_{\text{div}}(\Omega)$ be an arbitrary function. Multiplication of (7.129) with $\boldsymbol{v}$ and applying integration by parts gives

$$\begin{aligned} 0 &= \frac{\tau_q}{\Delta t}\left(\boldsymbol{u}_{n+1} - \tilde{\boldsymbol{u}}_{n+1}, \boldsymbol{v}\right) + \left(\nabla\varphi\left(\tilde{\boldsymbol{u}}_{n+1}\right) + \nabla\left(p - \hat{p}\right), \boldsymbol{v}\right) \\ &= \frac{\tau_q}{\Delta t}\left(\boldsymbol{u}_{n+1} - \tilde{\boldsymbol{u}}_{n+1}, \boldsymbol{v}\right) + \int_\Gamma \left(\varphi\left(\tilde{\boldsymbol{u}}_{n+1}\right) + p - \hat{p}\right) \boldsymbol{v}\cdot\boldsymbol{n}\, ds - \left(\nabla\cdot\boldsymbol{v}, \varphi\left(\tilde{\boldsymbol{u}}_{n+1}\right) + p - \hat{p}\right) \\ &= \frac{\tau_q}{\Delta t}\left(\boldsymbol{u}_{n+1} - \tilde{\boldsymbol{u}}_{n+1}, \boldsymbol{v}\right), \end{aligned}$$

since $\boldsymbol{v}\cdot\boldsymbol{n}$ vanishes on Γ, $\varphi\left(\tilde{\boldsymbol{u}}_{n+1}\right) + p - \hat{p} \in L^2(\Omega)$, and $\nabla\cdot\boldsymbol{v} = 0$ in $L^2(\Omega)$. It follows that

$$\left(\boldsymbol{u}_{n+1} - \tilde{\boldsymbol{u}}_{n+1}, \boldsymbol{v}\right) = 0 \quad \forall\ \boldsymbol{v} \in H_{\text{div}}(\Omega),\ n = 0, 1, 2, \ldots.$$

This relation is just the definition of the $L^2(\Omega)$ projection into $H_{\text{div}}(\Omega)$. The estimate of the norms follows directly from (3.164). ■

Example 7.84 (The Non-incremental Pressure-Correction Scheme) This scheme is the simplest pressure-correction scheme, proposed in Chorin (1968) and in Témam (1969). It is given by

$$\hat{p} = 0 \quad \text{in (7.127)}, \qquad \varphi\left(\tilde{\boldsymbol{u}}_{n+1}\right) = 0 \quad \text{in (7.129)}.$$

Numerical analysis shows that the magnitude of the splitting error of this scheme is of first order such that the use of a first order time stepping scheme is sufficient.

The non-incremental pressure-correction scheme with backward Euler time stepping ($q = 1$, $\tau_1 = 1$, $\tau_0 = -1$ in (7.126)) has the following form: Given $\boldsymbol{u}_0$, compute $(\tilde{\boldsymbol{u}}_{n+1}, \boldsymbol{u}_{n+1}, p_{n+1})$ by solving

$$\begin{aligned} \tilde{\boldsymbol{u}}_{n+1} + \Delta t_{n+1}\left(-\nu\Delta\tilde{\boldsymbol{u}}_{n+1} + (\boldsymbol{u}_n\cdot\nabla)\tilde{\boldsymbol{u}}_{n+1}\right) &= \boldsymbol{u}_n + \Delta t_{n+1}\boldsymbol{f}_{n+1} && \text{in } \Omega, \\ \tilde{\boldsymbol{u}}_{n+1} &= \boldsymbol{0} && \text{on } \Gamma, \end{aligned} \tag{7.130}$$

and

$$\begin{aligned} \boldsymbol{u}_{n+1} + \Delta t_{n+1}\nabla p_{n+1} &= \tilde{\boldsymbol{u}}_{n+1} && \text{in } \Omega, \\ \nabla\cdot\boldsymbol{u}_{n+1} &= 0 && \text{in } \Omega, \\ \boldsymbol{u}_{n+1}\cdot\boldsymbol{n} &= 0 && \text{on } \Gamma, \end{aligned} \tag{7.131}$$

for $n = 1, 2, 3, \ldots$.

To obtain a scalar equation for the pressure from (7.131), the (negative of the) divergence of the first equation of (7.131) is considered. This approach gives, using also the second equation,

$$-\nabla\cdot\nabla p_{n+1} = -\Delta p_{n+1} = -\frac{1}{\Delta t_{n+1}}\nabla\cdot\tilde{\boldsymbol{u}}_{n+1}. \tag{7.132}$$

Equation (7.132) is just a Poisson equation for the pressure, the so-called pressure-projection equation, which still has to be equipped with boundary conditions. From (7.131) and from the boundary conditions for $\tilde{\boldsymbol{u}}_{n+1}$ and $\boldsymbol{u}_{n+1}$, one obtains

$$\nabla p_{n+1}\cdot\boldsymbol{n} = -\frac{1}{\Delta t_{n+1}}\left(\boldsymbol{u}_{n+1}-\tilde{\boldsymbol{u}}_{n+1}\right)\cdot\boldsymbol{n} = 0. \tag{7.133}$$

Thus, homogeneous Neumann boundary conditions are applied.

After having computed p_{n+1}, the velocity $\boldsymbol{u}_{n+1}$ can be recovered from (7.131)

$$\boldsymbol{u}_{n+1} = \tilde{\boldsymbol{u}}_{n+1} - \Delta t_{n+1}\nabla p_{n+1}. \tag{7.134}$$

□

Remark 7.85 (Numerical Analysis for the Non-incremental Pressure-Correction Scheme) A linear convergence for the solution computed with the non-incremental pressure-correction scheme with respect to an equidistant length of the time step Δt was proved in Shen (1992a). The analysis required a regularity property of the temporal derivative of the pressure. This requirement is, however, generally not satisfied for general flows because of the lack of compatibility of the data as $t\to 0$, see Remark 7.30.

In Prohl (1997, Theorem 6.1) optimal error estimates were derived without the regularity assumption on the pressure. This analysis requires assumptions on the regularity of the velocity and of the right-hand side, in particular on temporal derivatives of the right-hand side. Provided that the Navier–Stokes equations (7.1) possess a unique solution in $(0,T]$, then for sufficiently small time steps the following estimates can be proved

$$\max_{1\le n\le N}\left(\|\boldsymbol{u}(t_n)-\tilde{\boldsymbol{u}}_n\|_{L^2(\Omega)} + \min\{t_n,1\}\,\|p(t_n)-p_n\|_{H^{-1}(\Omega)}\right) \le C\Delta t$$

and

$$\max_{1\le n\le N}\left(\|\nabla\left(\boldsymbol{u}(t_n)-\tilde{\boldsymbol{u}}_n\right)\|_{L^2(\Omega)} + \min\{t_n,1\}^{1/2}\,\|p(t_n)-p_n\|_{L^2(\Omega)}\right) \le C\,(\Delta t)^{1/2}, \tag{7.135}$$

where $T = N\Delta t$ and the constants depend only on the data of the problem. It is mentioned in Prohl (1997, Remark 6.1) that the same estimates can be proved for $\boldsymbol{u}_n$ instead of $\tilde{\boldsymbol{u}}_n$. □

Remark 7.86 (On the Non-incremental Pressure-Correction Scheme)

- If the convective terms $(\tilde{\boldsymbol{u}}_n \cdot \nabla)\tilde{\boldsymbol{u}}_{n+1}$ or $(\tilde{\boldsymbol{u}}_n \cdot \nabla)\tilde{\boldsymbol{u}}_n$ are used, the divergence-free velocity $\boldsymbol{u}_n$ appears only on the right-hand side of the momentum equation. Then, it can be replaced with (7.134) for t_n leading to the equation

$$\tilde{\boldsymbol{u}}_{n+1} + \Delta t_{n+1} \left(-\nu\Delta\tilde{\boldsymbol{u}}_{n+1} + (\tilde{\boldsymbol{u}}_n \cdot \nabla)\tilde{\boldsymbol{u}}_{n+1}\right) = \tilde{\boldsymbol{u}}_n - \Delta t_n \nabla p_n + \Delta t_{n+1}\boldsymbol{f}_{n+1}.$$

  Since also in the pressure-correction equation (7.133) only $\tilde{\boldsymbol{u}}_{n+1}$ is present, the divergence-free velocity does not appear in this formulation of the scheme. If it is not needed for some other reason, it is not necessary to compute the update (7.134).
- The artificial Neumann boundary condition for the pressure (7.133) induces a numerical boundary layer that prevents the scheme to obtain a first order convergence in (7.135), see Rannacher (1992). In Prohl (1997, Theorem 6.2), local estimates were proved that show that the pressure converges of first order in the interior of Ω, to be precise in $L^2(\Omega')$ with $\Omega' \subset\subset \Omega$.
- It can be shown that the scheme has a splitting error of order $\mathcal{O}(\Delta t)$ that cannot be reduced. Hence, using a higher order time stepping scheme does not lead to an improvement of the order of convergence.

□

Remark 7.87 (On the Stability of the Spatial Discretization) For the numerical solution, (7.130) and (7.132) has to be discretized in space. Since it is known that the solution of a Poisson problem with homogeneous Neumann boundary conditions is unique up to an additive constant, this method does not seem to need an inf-sup condition for obtaining a unique pressure. Hence, arbitrary finite element spaces can be used for velocity and pressure. Because of the large splitting error, the use of low order discretizations like P_1/P_1 is particularly attractive.

In fact, the pressure-correction equation (7.132) is kind of a strong form of the mass balance equation that appears in the PSPG stabilization, see Sects. 4.5.1 and 5.3.2. To illustrate this connection, assume that all functions in the Navier–Stokes equations are sufficiently smooth. Taking the divergence of the momentum equation yields

$$-\Delta p = -\nabla \cdot (\boldsymbol{f} - \partial_t \boldsymbol{u} + \nu\Delta\boldsymbol{u} - (\boldsymbol{u} \cdot \nabla)\boldsymbol{u}) .$$

Together with the mass balance, one obtains with some parameter $\delta^p > 0$

$$\nabla \cdot \boldsymbol{u} - \delta^p \Delta p = -\delta^p \nabla \cdot (\boldsymbol{f} - \partial_t \boldsymbol{u} + \nu\Delta\boldsymbol{u} - (\boldsymbol{u} \cdot \nabla)\boldsymbol{u}) .$$

Transferring this equation into a weak form gives with integration by parts and the boundary condition (7.133)

$$(\nabla \cdot \boldsymbol{u}, q) + \delta^p (\nabla p, \nabla q) = -\delta^p (\nabla \cdot (\boldsymbol{f} - \partial_t \boldsymbol{u} + \nu \Delta \boldsymbol{u} - (\boldsymbol{u} \cdot \nabla)\boldsymbol{u}), q) \qquad (7.136)$$

for all $q \in H_0^1(\Omega)$. This equation is of the same form as the PSPG stabilization, e.g., see (5.36) for the Oseen equations. The left-hand side of (7.136) is responsible for the stabilization with respect to the discrete inf-sup condition whereas the right-hand side of (7.136) accounts for the consistency of the method. Approximating this right-hand side by zero gives for the corresponding strong form

$$-\Delta p = -\frac{1}{\delta^p} \nabla \cdot \boldsymbol{u},$$

which is of the same form as the pressure-projection equation (7.132). In particular, one can see that the stabilization parameter is of size $\delta^p = \Delta t_{n+1}$. Thus, the stabilization becomes small for small time steps. In fact, it is shown in Guermond and Quartapelle (1998) that in the case of using finite element spaces for velocity and pressure that do not satisfy the discrete inf-sup condition (3.51) spurious oscillations may appear if the time step becomes too small, in particular if $\Delta t < Ch^2$ in the case that the pair of spaces P_1/P_1 is used. For small time steps, it is proposed to change the term Δt_{n+1} in (7.132) to $\Delta t_{n+1} + h^2$. The interpretation of the non-incremental pressure-projection equation as a kind of PSPG stabilization was first given in Rannacher (1992). □

Remark 7.88 (The Non-incremental Pressure-Correction Scheme with Grad-Div Stabilization) The non-incremental pressure-correction scheme with grad-div stabilization, compare Sect. 4.6.1, or sparse grad-div stabilization, see Remark 4.129, was proposed and analyzed in Bowers et al. (2014). Applying either type of grad-div stabilization improved the accuracy of the computed solutions in the numerical studies presented in this paper. However, both the grad-div stabilization and the sparse grad-div stabilization introduce couplings between different components of the velocity field, see Remark 4.129. Thus, instead of having to solve d scalar equations for the velocity, a coupled system has to be solved, which has a negative impact on the efficiency of the non-incremental pressure-correction scheme. □

Remark 7.89 (The Standard Incremental Pressure-Correction Scheme) The goal of this scheme is to increase the order of convergence compared with the non-incremental pressure-correction scheme by using a better approximation for $\hat{p}$. A natural choice is the pressure from the previous discrete time

$$\hat{p} = p_n \quad \text{in (7.127)}, \qquad \varphi(\tilde{\boldsymbol{u}}_{n+1}) = 0 \quad \text{in (7.129)}.$$

Since the scheme should perform better than first order in space, it should be combined with a higher order temporal discretization than the backward Euler scheme. Popular second order discretizations are the Crank–Nicolson scheme, see Example 7.49, and the backward difference formula 2 (BDF2) scheme, see Remark 7.64. The standard incremental pressure-correction scheme became popular with the paper (van Kan 1986). □

Example 7.90 (The Standard Incremental Pressure-Correction Scheme with the Crank–Nicolson Scheme) There are different proposals for this kind of scheme that differ in the form of the convective term, e.g., see Shen (1992b, 1996) and Prohl (1997, p. 141). These schemes do not fit exactly in the form (7.126) since the right-hand side is evaluated in the middle of the time interval and an average of velocities is used in the viscous term and in Shen (1996) also in the convective term.

The proposal from Shen (1996) has the form

$$\begin{aligned} \tilde{\boldsymbol{u}}_{n+1} + \Delta t_{n+1}\Big(-\nu\Delta\hat{\boldsymbol{u}}_{n+1/2} + \big(\hat{\boldsymbol{u}}_{n+1/2}\cdot\nabla\big)\,\hat{\boldsymbol{u}}_{n+1/2} \\ +\frac{1}{2}\left(\nabla\cdot\hat{\boldsymbol{u}}_{n+1/2}\right)\hat{\boldsymbol{u}}_{n+1/2} + \nabla p_n\Big) &= \boldsymbol{u}_n + \Delta t_{n+1}\boldsymbol{f}_{n+1/2} \quad \text{in } \Omega, \\ \tilde{\boldsymbol{u}}_{n+1} + \boldsymbol{u}_n &= \boldsymbol{0} \qquad\qquad\qquad \text{on } \Gamma, \end{aligned}$$

with $\hat{\boldsymbol{u}}_{n+1/2} = (\tilde{\boldsymbol{u}}_{n+1} + \boldsymbol{u}_n)/2$ and

$$\begin{aligned} \boldsymbol{u}_{n+1} + \Delta t_{n+1}\nabla\left(\frac{p_{n+1}-p_n}{2}\right) &= \tilde{\boldsymbol{u}}_{n+1} \text{ in } \Omega, \\ \nabla\cdot\boldsymbol{u}_{n+1} &= 0 \qquad \text{in } \Omega, \\ \boldsymbol{u}_{n+1}\cdot\boldsymbol{n} &= 0 \qquad \text{on } \Gamma, \end{aligned} \tag{7.137}$$

$n = 0, 1, 2, \ldots$. Note that the convective term is just the strong version of the divergence form, compare (6.20). The pressure-projection equation derived from (7.137) by taking the negative of the divergence has the form

$$-\Delta\left(p_{n+1} - p_n\right) = -\frac{2}{\Delta t_{n+1}}\nabla\cdot\tilde{\boldsymbol{u}}_{n+1} \quad \text{in } \Omega. \tag{7.138}$$

With the boundary conditions for $\boldsymbol{u}_{n+1}$ and $\tilde{\boldsymbol{u}}_{n+1}$, one obtains from (7.137) the boundary condition

$$\nabla\left(p_{n+1} - p_n\right)\cdot\boldsymbol{n} = 0 \quad \text{on } \Gamma \tag{7.139}$$

for (7.138). □

Remark 7.91 (Error Analysis of the Standard Incremental Pressure-Correction Scheme with Crank–Nicolson Scheme) Error estimates for the standard incremental pressure-correction scheme with Crank–Nicolson scheme can be found in Shen (1992b, 1996) and Prohl (1997, Chap. 7). The analysis in Prohl (1997) was actually performed for the time-dependent Stokes equations (7.125) but it is mentioned in Prohl (1997, Remark 7.1) that it can be extended to the instationary Navier–Stokes equations. It is noted also in Shen (1996) that the estimate of the convective term is technically involved but the main contribution of the error bound comes from the estimate of an auxiliary time-dependent Stokes problem.

The analysis in Shen (1992b), considers an implicit version of the convective term. It uses a compatibility assumption that is generally not satisfied.

In Shen (1996) a neighborhood of the initial time $t = 0$ is neglected in the analysis. Concretely, it is assumed that at a time $t_0 > 0$ there are approximations for velocity and pressure that are sufficiently accurate. With further assumptions on the velocity and the right-hand side, an error estimate of the form

$$\begin{aligned}\Delta t \sum_{n=1}^{N} \|\boldsymbol{u}(t_n) - \tilde{\boldsymbol{u}}_n\|_{L^2(\Omega)}^2 + (\Delta t)^2 \|\nabla (\boldsymbol{u}(t_N) - \tilde{\boldsymbol{u}}_N)\|_{L^2(\Omega)}^2 \\ + (\Delta t)^2 \|p(t_N) - p_n\|_{L^2(\Omega)}^2 \le C (\Delta t)^4, \quad 0 < t_0 \le t \le T,\end{aligned}$$

was proved, where N is the largest integer contained in $(t - t_0)/\Delta t$ and C depends on the data and on t_0.

The analysis in Prohl (1997) aimed for error estimates in different norms. With regularity assumptions on the solution of the continuous problem, error estimates for the gradient of the velocity

$$\max_{1\le n\le N} \left\| \nabla \left(\boldsymbol{u}(t_{n+1/2}) - \frac{\tilde{\boldsymbol{u}}_{n+1} + 2\tilde{\boldsymbol{u}}_n + \tilde{\boldsymbol{u}}_{n-1}}{4} \right) \right\|_{L^2(\Omega)} \le C \Delta t \log \left(\frac{1}{\Delta t} \right),$$

for the velocity

$$\max_{1\le n\le N} \left(\min\{t_{n+1}, 1\} \left\| \boldsymbol{u}(t_n) - \frac{\tilde{\boldsymbol{u}}_{n+1} + 2\tilde{\boldsymbol{u}}_n + \tilde{\boldsymbol{u}}_{n-1}}{4} \right\|_{L^2(\Omega)} \right) \le C (\Delta t)^2 \log \left(\tfrac{1}{\Delta t} \right),$$

and for the pressure

$$\max_{1\le n\le N} \left(\min\{t_{n+1/2}, 1\} \left\| p(t_{n+1/2}) - \frac{p_{n+1} + 2p_n + p_{n-1}}{4} \right\|_{L^2(\Omega)} \right) \le C \Delta t \log \left(\tfrac{1}{\Delta t} \right),$$

were obtained, see Prohl (1997, Theorem 7.1). These bounds are valid also for the divergence-free velocity.

In summary, the convergence of the standard incremental pressure-correction scheme with Crank–Nicolson scheme is of higher order with respect to Δt than the convergence of the non-incremental pressure-correction scheme, compare (7.135).

□

Example 7.92 (The Standard Incremental Pressure-Correction Scheme with the BDF2 Scheme) The standard incremental pressure-correction scheme with BDF2 ($q = 2$, $\tau_2 = 3/2$, $\tau_1 = -2$, $\tau_0 = 1/2$ in (7.126)) and equidistant time step Δt has the form: Given $(\boldsymbol{u}_0, p_0)$ and $(\boldsymbol{u}_1, p_1)$, compute $(\tilde{\boldsymbol{u}}_{n+1}, \boldsymbol{u}_{n+1}, p_{n+1})$ by solving

$$\begin{aligned}3\tilde{\boldsymbol{u}}_{n+1} &+ 2\Delta t\left(-\nu\Delta\tilde{\boldsymbol{u}}_{n+1} + (\boldsymbol{u}_n\cdot\nabla)\tilde{\boldsymbol{u}}_{n+1}\right)\\ &= 4\boldsymbol{u}_n - \boldsymbol{u}_{n-1} + 2\Delta t\left(\boldsymbol{f}_{n+1} - \nabla p_n\right) \quad \text{in } \Omega,\\ \tilde{\boldsymbol{u}}_{n+1} &= \boldsymbol{0} \quad \text{on } \Gamma,\end{aligned} \tag{7.140}$$

and

$$\begin{aligned}3\boldsymbol{u}_{n+1} + 2\Delta t\nabla\left(p_{n+1} - p_n\right) &= 3\tilde{\boldsymbol{u}}_{n+1} && \text{in } \Omega,\\ \nabla\cdot\boldsymbol{u}_{n+1} &= 0 && \text{in } \Omega,\\ \boldsymbol{u}_{n+1}\cdot\boldsymbol{n} &= 0 && \text{on } \Gamma,\end{aligned} \tag{7.141}$$

for $n = 1, 2, 3, \ldots$. Note that modifications of the scheme are possible by using other forms of the convective term, in particular extrapolated forms like (7.128) can be found in the literature. By applying the negative of the divergence to (7.141), one obtains the following equation for the update of the pressure

$$-\Delta\left(p_{n+1} - p_n\right) = -\frac{3}{2\Delta t}\nabla\cdot\tilde{\boldsymbol{u}}_{n+1} \quad \text{in } \Omega, \tag{7.142}$$

which is equipped with the same boundary condition (7.139) as in the case of the Crank–Nicolson scheme. □

Remark 7.93 (The Initial Step) Besides the initial velocity field $\boldsymbol{u}_0$, the standard incremental pressure-correction scheme with BDF2 requires the computation of quantities $p_0 = p(0, \boldsymbol{x})$ and $(\boldsymbol{u}_1, p_1)$. An initial pressure can be computed as described in Remark 7.29. Then, $(\boldsymbol{u}_1, p_1)$ can be computed with the incremental pressure-correction scheme, where a first order one-step discretization in time is applied

$$\begin{aligned}\tilde{\boldsymbol{u}}_1 + \Delta t\left(-\nu\Delta\tilde{\boldsymbol{u}}_1 + (\boldsymbol{u}_0\cdot\nabla)\tilde{\boldsymbol{u}}_1\right) &= \boldsymbol{u}_0 + \Delta t\left(\boldsymbol{f}_1 - \nabla p_0\right) && \text{in } \Omega,\\ \tilde{\boldsymbol{u}}_1 &= \boldsymbol{0} && \text{on } \Gamma,\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{u}_1 + \Delta t\nabla\left(p_1 - p_0\right) &= \tilde{\boldsymbol{u}}_1 && \text{in } \Omega,\\ \nabla\cdot\boldsymbol{u}_1 &= 0 && \text{in } \Omega,\\ \boldsymbol{u}_1\cdot\boldsymbol{n} &= 0 && \text{on } \Gamma.\end{aligned}$$

It can be proved that this step is sufficiently accurate such that the order of convergence of the complete scheme is not spoiled, see Guermond and Shen (2004). □

Remark 7.94 (Error Analysis of the Standard Incremental Pressure-Correction Scheme with BDF2 Scheme) The fully discretized standard incremental pressure-correction scheme with BDF2 and with the Galerkin finite element method was analyzed in Guermond (1999b). An explicit form of the convective term was considered that is based on an extrapolation from the previous two times, like (7.128). If the solution of the continuous problem is sufficiently regular, then estimates of the form

$$\max_{1\le n\le N} \|\boldsymbol{u}(t_n) - \tilde{\boldsymbol{u}}_n\|_{H^1(\Omega)} + \max_{1\le n\le N} \|p(t_n) - p_n\|_{L^2(\Omega)} \le C\Delta t, \tag{7.143}$$

$$(\Delta t)^{1/2} \sum_{n=1}^{N} \|\boldsymbol{u}(t_n) - \tilde{\boldsymbol{u}}_n\|_{L^2(\Omega)} \le C\Delta t^2$$

were proved. The constants depend on data of the problem and norms of the solution. □

Remark 7.95 (On the Standard Incremental Pressure-Correction Scheme)

- The boundary condition (7.139) implies

$$\nabla p_{n+1} \cdot \boldsymbol{n} = \nabla p_n \cdot \boldsymbol{n} = \ldots = \nabla p_0 \cdot \boldsymbol{n} \text{ on } \Gamma.$$

  This boundary condition is an unphysical boundary condition that leads to numerical boundary layers.
- The scheme has an irreducible splitting error of order $\mathcal{O}\left(\Delta t^2\right)$. Hence, the use of a higher than second order discretization in time does not improve the order of convergence.

□

Remark 7.96 (Stability of the Spatial Discretization) The pressure-correction equations (7.138) and (7.142) can be written in the form

$$-\nabla \cdot \tilde{\boldsymbol{u}}_{n+1} + C\Delta t\Delta\,(p_{n+1} - p_n) = 0 \quad \text{in } \Omega,$$

with $C = 1/2$ for the Crank–Nicolson scheme and $C = 2/3$ for BDF2. Hence, there is some pressure contribution in the continuity equation that acts as stabilization with respect to the discrete inf-sup condition (3.51) if the difference $p_{n+1} - p_n$ is sufficiently large. Otherwise, in particular if a steady-state solution is approached, the stabilization vanishes. In fact, it is reported in the literature, e.g., in Guermond and Quartapelle (1998) and de Frutos et al. (2016c), that solutions with spurious oscillations are obtained in this case if spatial discretizations are used that do not satisfy the discrete inf-sup condition.

A remedy consists in introducing an additional stabilization of PSPG-type in the pressure-correction equation, leading to an equation of the form

$$-\Delta\left(\left(1+\frac{\delta^p}{C\Delta t}\right)p_{n+1}-p_n\right) = -\frac{1}{C\Delta t}\nabla\cdot\tilde{\boldsymbol{u}}_{n+1} \text{ in } \Omega,$$
$$\nabla\left(\left(1+\frac{\delta^p}{C\Delta t}\right)p_{n+1}-p_n\right)\cdot\boldsymbol{n} = 0 \qquad \text{on } \Gamma.$$

A modification of this type was proposed in Prohl (1997, Chap. 7.4) with $\delta^p = h^2$. Generally, one can choose the stabilization parameter locally in the form $\delta_K^p = \delta_0 h_K^2$ with some positive constant δ_0, see Sect. 4.5.1.

Another possibility for introducing a stabilization, which was investigated in Codina (2001a), consists in using fluctuations of the pressure gradient, compare also Remark 4.111. □

Example 7.97 (The Rotational Incremental Pressure-Correction Scheme) The goal of this scheme consists in curing the inaccuracies caused by the artificial boundary condition of the pressure in the standard incremental pressure-correction scheme. This scheme was proposed in Timmermans et al. (1996). It is obtained by choosing

$$\hat{p} = p_n \quad \text{in (7.127)}, \qquad \varphi\left(\tilde{\boldsymbol{u}}_{n+1}\right) = \nu\nabla\cdot\tilde{\boldsymbol{u}}_{n+1} \quad \text{in (7.129)}.$$

Using BDF2, the velocity step of this scheme is identical to (7.140) and the projection step has the form

$$\begin{aligned} 3\boldsymbol{u}_{n+1} + 2\Delta t\nabla\left(p_{n+1}-p_n\right) &= 3\tilde{\boldsymbol{u}}_{n+1} - 2\nu\Delta t\nabla\left(\nabla\cdot\tilde{\boldsymbol{u}}_{n+1}\right) && \text{in } \Omega, \\ \nabla\cdot\boldsymbol{u}_{n+1} &= 0 && \text{in } \Omega, \\ \boldsymbol{u}_{n+1}\cdot\boldsymbol{n} &= 0 && \text{on } \Gamma. \end{aligned} \tag{7.144}$$

The problem for the pressure becomes, using $\nabla\cdot\nabla = \Delta$,

$$-\Delta\left(p_{n+1}-p_n\right) = -\frac{3}{2\Delta t}\nabla\cdot\tilde{\boldsymbol{u}}_{n+1} + \nu\Delta\left(\nabla\cdot\tilde{\boldsymbol{u}}_{n+1}\right) \tag{7.145}$$

or

$$-\Delta\tilde{p}_n = -\frac{3}{2\Delta t}\nabla\cdot\tilde{\boldsymbol{u}}_{n+1} \quad \text{with} \quad \tilde{p}_n = p_{n+1}-p_n+\nu\nabla\cdot\tilde{\boldsymbol{u}}_{n+1}. \tag{7.146}$$

The boundary conditions for (7.145) and (7.146) will be discussed in Remark 7.98, see (7.148) for the result.

Using (7.146), the update for the pressure has to be computed by subtracting $\nu\nabla\cdot\tilde{\boldsymbol{u}}_{n+1}$ from $\tilde{p}_n$. If standard finite elements are used for velocity and pressure with a continuous pressure space, one has to modify (7.146) since the term $\nu\nabla\cdot\tilde{\boldsymbol{u}}_{n+1}$ is usually discontinuous. □

Remark 7.98 (On the Rotational Incremental Pressure-Correction Scheme) Inserting (7.144) in (7.140) yields

$$3\boldsymbol{u}_{n+1} + 2\Delta t\Big(\nu\nabla\left(\nabla\cdot\tilde{\boldsymbol{u}}_{n+1}\right) - \nu\Delta\tilde{\boldsymbol{u}}_{n+1} + (\boldsymbol{u}_n\cdot\nabla)\tilde{\boldsymbol{u}}_{n+1} + \nabla p_{n+1}\Big)$$
$$= 4\boldsymbol{u}_n - \boldsymbol{u}_{n-1} + 2\Delta t\boldsymbol{f}_{n+1}.$$

Using now (3.155) leads to

$$3\boldsymbol{u}_{n+1} + 2\Delta t\Big(\nu\nabla\times\nabla\times\tilde{\boldsymbol{u}}_{n+1} + (\boldsymbol{u}_n\cdot\nabla)\tilde{\boldsymbol{u}}_{n+1} + \nabla p_{n+1}\Big)$$
$$= 4\boldsymbol{u}_n - \boldsymbol{u}_{n-1} + 2\Delta t\boldsymbol{f}_{n+1}.$$

From (7.144), it follows with (3.156) that $\nabla\times\tilde{\boldsymbol{u}}_{n+1} = \nabla\times\boldsymbol{u}_{n+1}$ such that an equation of the following form is obtained

$$\begin{aligned}
3\boldsymbol{u}_{n+1} + 2\Delta t\left(\nu\nabla\times\nabla\times\boldsymbol{u}_{n+1} + (\boldsymbol{u}_n\cdot\nabla)\tilde{\boldsymbol{u}}_{n+1} + \nabla p_{n+1}\right)& \\
= 4\boldsymbol{u}_n - \boldsymbol{u}_{n-1} + 2\Delta t\boldsymbol{f}_{n+1} &\quad \text{in } \Omega,\\
\nabla\cdot\boldsymbol{u}_{n+1} = 0 &\quad \text{in } \Omega,\\
\boldsymbol{u}_{n+1}\cdot\boldsymbol{n} = 0 &\quad \text{on } \Gamma.
\end{aligned} \tag{7.147}$$

Because of this form, the scheme is called rotational incremental pressure-correction scheme. From (7.147), also the boundary condition for the pressure can be derived by using the boundary conditions for $\tilde{\boldsymbol{u}}_{n+1}$, $\boldsymbol{u}_{n+1}$, $\boldsymbol{u}_n$, and $\boldsymbol{u}_{n-1}$

$$\nabla p_{n+1}\cdot\boldsymbol{n} = \left(\boldsymbol{f}_{n+1} - \nu\nabla\times\nabla\times\boldsymbol{u}_{n+1}\right)\cdot\boldsymbol{n} \text{ on } \Gamma. \tag{7.148}$$

Considering now the Navier–Stokes equations (7.1), discretizing this equation in time with BDF2, denoting the solution at time t_n with $(\boldsymbol{u}_n, p_n)$, and using (3.155) gives for the momentum balance

$$3\boldsymbol{u}_{n+1} + 2\Delta t\Big(\nu\nabla\times\nabla\times\boldsymbol{u}_{n+1} - \nu\nabla\left(\nabla\cdot\boldsymbol{u}_{n+1}\right) + (\boldsymbol{u}_{n+1}\cdot\nabla)\boldsymbol{u}_{n+1} + \nabla p_{n+1}\Big)$$
$$= 4\boldsymbol{u}_n - \boldsymbol{u}_{n-1} + 2\Delta t\boldsymbol{f}_{n+1}.$$

Restricting this equation to the boundary and multiplying it with $\boldsymbol{n}$ yields exactly the boundary condition (7.148) since the velocity field in the Navier–Stokes equations is divergence-free and it vanishes on the boundary. Thus, the boundary condition (7.148) is a consistent boundary condition because it is not introduced by some numerical method but it can be derived directly from the Navier–Stokes equations. □

Remark 7.99 (Error Analysis of the Rotational Incremental Pressure-Correction Scheme) An error analysis of the rotational incremental pressure-correction scheme was performed in Guermond and Shen (2004). The time-dependent Stokes equations (7.125) were considered to concentrate in the analysis on the splitting. BDF2 was used as temporal discretization, with the initial step as described in Remark 7.93. If the solution of the continuous problem is sufficiently smooth in time and space, then the following error estimates were proved

$$\Delta t \sum_{n=0}^{N} \|\boldsymbol{u}(t_n) - \boldsymbol{u}_n\|_{L^2(\Omega)}^2 + \Delta t \sum_{n=0}^{N} \|\boldsymbol{u}(t_n) - \tilde{\boldsymbol{u}}_n\|_{L^2(\Omega)}^2 \le C\,(\Delta t)^4$$

and

$$\Delta t \sum_{n=0}^{N} \|\nabla\,(\boldsymbol{u}(t_n) - \boldsymbol{u}_n)\|_{L^2(\Omega)}^2 + \Delta t \sum_{n=0}^{N} \|\nabla\,(\boldsymbol{u}(t_n) - \tilde{\boldsymbol{u}}_n)\|_{L^2(\Omega)}^2$$
$$+ \Delta t \sum_{n=0}^{N} \|p(t_n) - p_n\|_{L^2(\Omega)}^2 \le C\,(\Delta t)^3\,,$$

with the constants depending on norms of the solution and on the final time. Note that the order of convergence for the gradient of the velocity and the pressure is 1.5, in contrast to the first order convergence of the standard incremental pressure-correction scheme, see (7.143). □

Remark 7.100 (On $\boldsymbol{u}_{n+1}$ and $\tilde{\boldsymbol{u}}_{n+1}$)

- In implementing and using pressure-correction schemes, one has two velocity fields in the computation: $\boldsymbol{u}_{n+1}$ and $\tilde{\boldsymbol{u}}_{n+1}$. The field $\boldsymbol{u}_{n+1}$ is (discretely) divergence-free but it does not satisfy the no-slip boundary condition. Only the normal component of $\boldsymbol{u}_{n+1}$ is required to vanish at the boundary, but not the tangential component(s). On the contrary, the boundary condition is satisfied from $\tilde{\boldsymbol{u}}_{n+1}$, but this field is not (discretely) divergence-free. Thus, none of the fields is perfect.
- From the point of view of convergence orders, one obtains the same results for both, $\boldsymbol{u}_{n+1}$ and $\tilde{\boldsymbol{u}}_{n+1}$, e.g., see Remarks 7.85, 7.91, and 7.99.
- From the point of view of implementation, $\boldsymbol{u}_{n+1}$ has another disadvantage. After having computed the pressure by solving the Poisson equation (7.132), (7.138), (7.142), or (7.145) using standard continuous finite elements, then $\boldsymbol{u}_{n+1}$ can be recovered by (7.131), (7.137), (7.141), or (7.144). This recovery requires the computation of the gradient of the computed pressure, which will give in general a discontinuous finite element function. Consequently, $\boldsymbol{u}_{n+1}$ will be a discontinuous finite element function, too.
- Note that $\boldsymbol{u}_{n+1}$ is not needed in computing the pressure p_{n+1}.
- Using a formulation of the convective term which does not include the divergence-free velocity field, then it is usually easy to eliminate this field from the momentum equation. The idea consists in solving the equation for the

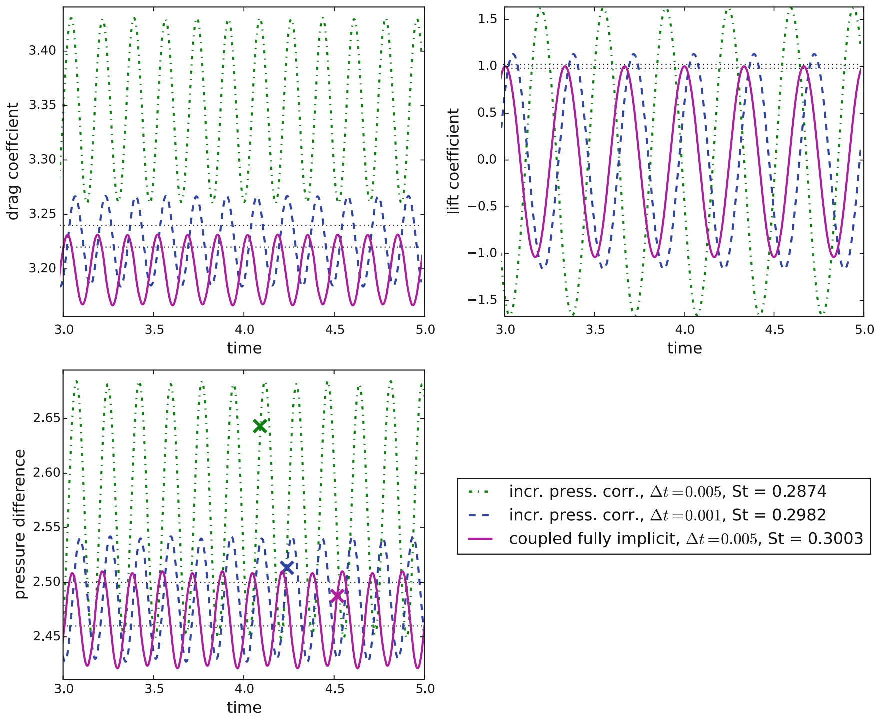

Fig. 7.11 Example 7.101. Quantities of interest. Reference intervals, see Table D.2, are depicted with the *dotted lines*. The *cross marks* the middle of the period. The reference interval for the Strouhal number is [0.295, 0.305]

projection for $\boldsymbol{u}_{n+1}$ and in inserting the result in the equation for the momentum balance, see Remark 7.86.

In the literature, e.g., in Guermond et al. (2006), it is advised to use only $\tilde{\boldsymbol{u}}_{n+1}$ in the implementation. □

Example 7.101 (Standard Incremental Pressure-Correction Scheme with BDF2 Scheme) Numerical results obtained with standard incremental pressure-correction scheme with BDF2 scheme (7.140)–(7.142) for a two-dimensional flow around a cylinder, Example D.8, are presented in Fig. 7.11. The simulations were performed with the Q_2/Q_1 Taylor–Hood pair of finite element spaces. The initial grid (level 0) is depicted in Fig. 6.5 and the results from Fig. 7.11 were computed on level 2 with 27,232 velocity degrees of freedom (including Dirichlet nodes) and 3480 pressure degrees of freedom. As comparison, results with the fully implicit method with Crank–Nicolson scheme (7.80) are presented.

First of all, it is noted that the pictures of the results on level 3 and 4 are very similar to the pictures from Fig. 7.11. Thus, it can be observed that the time step $\Delta t = 0.005$ suffices for the fully implicit approach to compute a solution where all quantities of interest are within their reference interval. The same time step gives for the incremental pressure-correction scheme very inaccurate results. Decreasing the length of the time step improves the accuracy considerably. But even for $\Delta t = 0.001$, most of the quantities of interest are not yet in their reference interval. Altogether, one can observe in this example that for the incremental pressure-correction scheme the temporal error is the dominant error. □

Chapter 8
The Time-Dependent Navier–Stokes Equations: Turbulent Flows

Remark 8.1 (Turbulent Flows) Usually, the behavior of incompressible turbulent flows is modeled with the incompressible Navier–Stokes equations (2.25). There is no mathematical definition of what is "turbulence". From the mathematical point of view, turbulent flows occur at high Reynolds numbers. From the physical point of view, these flows are characterized by possessing flow structures (eddies, scales) of very different sizes. Consider, e.g., a tornado. This tornado has some very large flow structures (large eddies) but also millions of very small flow structures.

The small scales are important for the physics of turbulent flows. A numerical scheme that simply neglects them, e.g., by introducing a large amount of artificial viscosity, computes a solution which is laminar and usually has much different properties than the turbulent solution. A prominent example is the mean velocity profile of a channel flow: it is of parabolic form for a laminar flow and it tends to become of a step profile for turbulent flows, see Fig. D.12. The way to treat the small scales, which cannot be resolved in simulations, consists in modeling their influence onto the resolved scales. With other words, a turbulence model has to be applied. A turbulence model has to contain less scales than the Navier–Stokes equations, which means, it has to be less complex than the Navier–Stokes equations. □

Remark 8.2 (Contents) This chapter presents some approaches for turbulence modeling. The emphasis will be on models that allow a mathematical or numerical analysis or whose derivation is based on mathematical arguments. Some remarks concerning the application of these models in practical simulations are given. However, the use of turbulence models in practice is a wide field of research and it is not the goal of this monograph to give a comprehensive presentation here. There are several monographs on this topic, e.g., Sagaut (2006), Lesieur et al. (2005) and Ferziger and Perić (1999).

With respect to the presentation of mathematical and numerical analysis there will be the emphasis on results that are obtained for the case of bounded domains. One can find in the literature a number of results for the space-periodic case. As already mentioned in Remark 2.31, the space-periodic case mimics the situation of

V. John, *Finite Element Methods for Incompressible Flow Problems*, Springer Series in Computational Mathematics 51, DOI 10.1007/978-3-319-45750-5_8

a domain without boundary and the periodic boundary conditions do not possess a physical meaning. From the analytical point of view, the absence of a boundary might simplify the analysis considerably and a somewhat different mathematical setup is used than in the case of a bounded domain. However, the interaction of a turbulent flow with the boundary of the domain is often of utmost importance in practice, such that results obtained for the case of a bounded domain are of more interest. □

8.1 Some Physical and Mathematical Characteristics of Turbulent Incompressible Flows

Remark 8.3 (Monographs) The physics of turbulent flows is the topic of a number of monographs, e.g., Pope (2000) and Davidson (2004). Mathematical aspects of turbulence flows are studied in Foias et al. (2001) and Chacón Rebollo and Lewandowski (2014). □

Remark 8.4 (The Incompressible Navier–Stokes Equations) For describing the physical properties of turbulent flows, it is sometimes convenient to use the incompressible Navier–Stokes equations with the Reynolds number in the viscous term

$$\begin{aligned} \partial_t \boldsymbol{u} - 2\mathrm{Re}^{-1}\nabla\cdot\mathbb{D}(\boldsymbol{u}) + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p &= \boldsymbol{f} \quad \text{in } (0,T]\times\Omega, \\ \nabla\cdot\boldsymbol{u} &= 0 \quad \text{in } (0,T]\times\Omega, \\ \boldsymbol{u}(0,\cdot) &= \boldsymbol{u}_0 \text{ in } \Omega. \end{aligned} \tag{8.1}$$

If $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, is a bounded domain, (8.1) has to be equipped with boundary conditions.

Turbulent flows are characterized by a high Reynolds number. In applications, the range of the Reynolds number for flows of this type starts around several thousand. Often, it is even larger by some orders of magnitude. In the case of high Reynolds numbers, the stabilizing forces in the momentum balance (the viscous term $2\mathrm{Re}^{-1}\nabla\cdot\mathbb{D}(\boldsymbol{u})$) are small compared with the destabilizing forces (the convective term $\partial_t\boldsymbol{u} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u}$, right-hand side of Newton's second law of motion (2.5)). □

Remark 8.5 (Instantaneous Flow and Statistics) Considering turbulent flows, one has the idea that the flow behaves in some sense chaotic and hardly predictable. In fact, small changes of the data might lead to a large change in the instantaneous flow field, i.e., in the flow field at a certain time $t \in (0,T]$ and in a certain point $\boldsymbol{x} \in \Omega$. In practice, one has never complete information about the data, e.g., on the initial condition in the whole domain Ω or on the boundary conditions at the complete boundary for all times. For instance, boundary conditions are available from measurement data only at some points at the boundary and at some times. Then, the needed boundary conditions for the equations are defined by an interpolation. Since the boundary conditions obtained in this way most probably differ from the actual boundary conditions, it is to be expected that the computed instantaneous flow field is also different from the actual one. From this point of view, the consideration of instantaneous turbulent flow fields is not of practical interest. Instead, one is interested in practice in statistics of the flow fields, often defined by averages in space and time. For different problems, different statistics are of importance, e.g., compare Examples D.12 and D.13. □

Remark 8.6 (On the Concept of Isotropic Turbulence) Much of the physical turbulence theory, e.g., the determination of the size of the smallest scales, is based on the concept of isotropic turbulence. A field $u(t,x)$ is called statistically stationary if all statistics of $u(t,x)$ are invariant under a shift of time. It is called statistically homogeneous if all statistics are invariant under a shift of position. If the field is also statistically invariant under rotations and reflections of the coordinate system, it is called (statistically) isotropic.

Wind tunnel experiments have been performed on (approximately) isotropic turbulence. However, isotropic turbulence is in general an idealization. □

Remark 8.7 (The Richardson Energy Cascade) Let $\Omega \subset \mathbb{R}^3$. In Richardson (1922), a description of the physical mechanisms was given that work in turbulent flows. Kinetic energy enters the flow at the largest eddies. Large eddies are unstable and break up into smaller ones. Thereby energy is transferred in the mean from the eddies of a given size to the next smaller eddies. This transfer in the mean does not exclude a local (in time and space) transfer in the opposite direction from smaller to larger eddies, mainly the next larger eddies, the so-called backscatter of energy. The smaller eddies undergo a similar process like the large eddies. This process is continued until the Reynolds number $\mathrm{Re}(l) = u(l)l/\nu$ of the eddies of size l is sufficiently small (of order one) such that the eddy motion is stable and the molecular viscosity becomes effective in dissipating the kinetic energy. This process is called energy cascade.

The size of the smallest eddies will be denoted by λ. □

Remark 8.8 (The Rate of Dissipation of Turbulent Energy) Denote by ε $[\mathrm{m}^2/\mathrm{s}^3]$ the rate of dissipation of turbulent energy which is defined in the following way. Consider $\boldsymbol{u}$ as a random variable and let $\langle \boldsymbol{u} \rangle$ be the mean value (expectation) of $\boldsymbol{u}$. The difference $\boldsymbol{u}' := \boldsymbol{u} - \langle \boldsymbol{u} \rangle$ is called the fluctuations. The rate of dissipation of turbulence energy is now defined by

$$\varepsilon := 2\nu \langle \mathbb{D}\left(\boldsymbol{u}'\right) : \mathbb{D}\left(\boldsymbol{u}'\right) \rangle.$$

The detailed theoretical and experimental study of particular flows shows that

$$\varepsilon = \mathcal{O}\left(\frac{U^3}{L}\right) \tag{8.2}$$

independently of Re, Pope (2000, p. 183). □

Remark 8.9 (The Kolmogorov Hypotheses) In the fundamental paper (Kolmogoroff 1941) three hypotheses about turbulent flows were postulated:

1. At sufficiently high Reynolds numbers, the small scale turbulent motions are isotropic.
2. In every turbulent flow at sufficiently high Reynolds number, the statistics of the small scale motions have a universal form that is uniquely given by ν and ε.
3. In every turbulent flow at sufficiently high Reynolds number, the statistics of motions of scales of size l in the range $L \gg l \gg \lambda$ have a universal form uniquely determined by ε and independent of ν.

□

Remark 8.10 (The Size of the Kolmogorov Scales) For describing the size of the smallest scales, the first and second hypothesis are of importance.

Let ε and ν be given. Then, there are unique length, velocity, and time scales which can be defined, the so-called Kolmogorov scales

$$\lambda = \left(\frac{\nu^3}{\varepsilon}\right)^{1/4} \text{[m]}, \quad u_\lambda = (\varepsilon\nu)^{1/4} \text{[m/s]}, \quad t_\lambda = \left(\frac{\nu}{\varepsilon}\right)^{1/2} \text{[s]}. \tag{8.3}$$

The scale λ is also called Kolmogorov length scale. The Reynolds number of eddies of size λ is

$$\text{Re}(\lambda) = \frac{\lambda u_\lambda}{\nu} = \frac{\nu\varepsilon^{1/4}}{\nu\varepsilon^{1/4}} = 1, \tag{8.4}$$

such that it is sufficiently small for the dissipation to be effective. The motion of eddies with $\text{Re}(\lambda) = 1$ is isotropic and the first hypothesis is met.

In addition, using (8.3), the rate of dissipation is given by the following expressions

$$\varepsilon^{1/2} = \frac{u_\lambda^2}{\nu^{1/2}}, \ \varepsilon^{1/2} = \frac{\nu^{3/2}}{\lambda^2} \quad \Longrightarrow \quad \varepsilon = \nu\frac{u_\lambda^2}{\lambda^2} \text{ and } \varepsilon = \nu\frac{1}{t_\lambda^2}, \tag{8.5}$$

such that

$$\frac{u_\lambda}{\lambda} = \frac{1}{t_\lambda} = \left(\frac{\varepsilon}{\nu}\right)^{1/2} \quad \text{[1/s]}. \tag{8.6}$$

The left-hand side is an approximation to the spatial derivative of the Kolmogorov velocity, which describes the velocity gradient, since λ is small. For the large eddies in turbulent flows, the velocity gradient increases with the Reynolds number since the flow field varies rapidly in space and time. Equation (8.6) shows that for the Kolmogorov scales, the velocity gradient is bounded uniformly with respect to the Reynolds number. It depends only on ν and ε. This dependency is required by the second Kolmogorov hypothesis. Altogether, (8.4) and (8.6) characterize the Kolmogorov scales as dissipative scales.

Now, one can estimate the size of the Kolmogorov scales. Using (8.2) and (8.3) gives

$$\frac{\lambda}{L} \sim \left(\frac{\nu^3}{L^3U^3}\right)^{1/4} \sim \mathrm{Re}^{-3/4} \quad \Longleftrightarrow \quad \lambda \sim \mathrm{Re}^{-3/4}, \tag{8.7}$$

where $L \sim 1$ m was assumed. □

Remark 8.11 (On the Impact of the Size of the Kolmogorov Scales in Numerical Simulations) A standard discretization of the Navier–Stokes equations (8.1), like the Galerkin finite element method studied in Sect. 7.2, seeks to simulate the behavior of all scales, including the Kolmogorov scales.

Consider as example the domain $\Omega = (0,1)^3$, such that $L = 1$, and a mesh of roughly 10^8 cubic mesh cells ($\approx 464^3$). Assuming that the mesh width is equal to the resolution of the discretization, as for low order finite elements, then scales of size $1/464$ can be represented on this mesh. Only those scales can be simulated which can be represented. Hence, a necessary condition for flows to be simulated on this grid is that for its Kolmogorov length it holds $\lambda \gtrapprox 1/464$. Assuming additionally that equality holds in (8.2), then it follows from (8.7) that flows up to a Reynolds number of $\mathrm{Re} \approx 464^{4/3} \approx 3590$ can be simulated. This is far less than the Reynolds number of turbulent flows in most applications.

Arguing the same way for a general situation, one finds with (8.7) that the number of degrees of freedom to resolve the Kolmogorov scales behaves like

$$\left(\frac{L}{\lambda}\right)^3 \sim \mathrm{Re}^{9/4}.$$

The application of the Galerkin finite element method (or any other standard discretization) is called Direct Numerical Simulation (DNS) in the context of turbulent flow simulations. The considerations in this remark show that a DNS is generally not feasible for the simulation of turbulent flows and it will not be feasible in the foreseeable future. □

Remark 8.12 (The Kinetic Energy Spectrum) The energy cascade, Remark 8.7, and the third hypothesis of Kolmogorov, Remark 8.9, can be expressed with the so-called kinetic energy spectrum.

Consider $\Omega = \mathbb{R}^3$ and a velocity field $\boldsymbol{u}(t,\boldsymbol{x})$. Applying the Fourier transform (A.23) gives

$$\mathcal{F}(\boldsymbol{u})(t,\boldsymbol{k}) = \int_{\mathbb{R}^3} \boldsymbol{u}(t,\boldsymbol{x}) e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\, d\boldsymbol{x},$$

where $\boldsymbol{k} = (k_1, k_2, k_3)^T$ is the wave number vector and the Fourier space is called wave number space. The kinetic energy for the wave number $\boldsymbol{k}$ is given by

$$E(t, \boldsymbol{k}) = \frac{1}{2}\langle \mathcal{F}(\boldsymbol{u})(t, \boldsymbol{k}) \cdot \mathcal{F}(\boldsymbol{u})^*(t, \boldsymbol{k})\rangle,$$

with $\mathcal{F}(\boldsymbol{u})^*$ denoting the conjugate complex of $\mathcal{F}(\boldsymbol{u})$ and $\langle \cdot, \cdot, \rangle$ denotes an appropriate mean value. Here, homogeneous isotropic turbulence with mean velocity equal to zero will be considered. For wave numbers with the same absolute value $k = \|\boldsymbol{k}\|_2$, one defines

$$E(t, k) = \int_{\|\boldsymbol{k}\|_2 = k} E(t, \boldsymbol{k})\, d\boldsymbol{k}.$$

Kolmogorov's third hypothesis implies a universal form of scales of size l with $L \gg l \gg \lambda$.

The kinetic energy spectrum of turbulent flows is sketched in Fig. 8.1, see Pope (2000, p. 229) for a detailed derivation. In the so-called inertial subrange, one finds that

$$E(k) = C_K \varepsilon^{2/3} k^{-5/3},$$

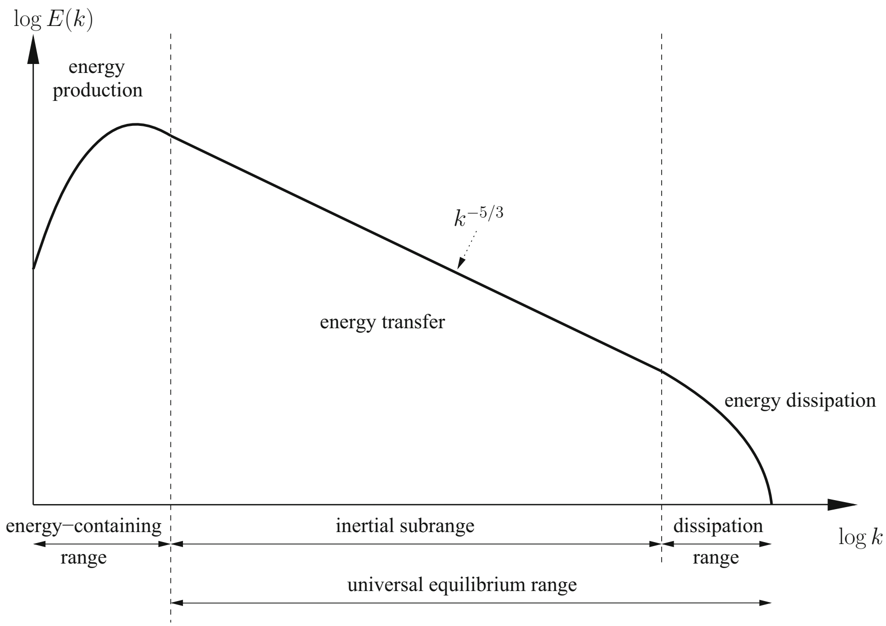

Fig. 8.1 Sketch of the kinetic energy spectrum of turbulent flows

where C_K is called Kolmogorov constant. This relation is often called Kolmogorov $-5/3$ spectrum.

Flows at high Reynolds number show a distinct inertial subrange which is absent for laminar flows. □

Remark 8.13 (Boundary Layers) Consider an example like a turbulent channel flow, see Example D.12. Then, there is a planar solid wall as boundary which is situated at $y = 0$, the domain at this boundary is given for $y > 0$, and a no-slip condition $\boldsymbol{u} = \boldsymbol{0}$ is applied at this wall. In this situation, there is a layer at the boundary. For the mean velocity field it holds that $\langle u_1 \rangle = \langle u_1 \rangle(y)$, $\langle u_2 \rangle = \langle u_3 \rangle = 0$. Then, the viscous stress (2.17) for the mean velocity is given by

$$\rho\nu \frac{\nabla\langle \boldsymbol{u} \rangle + (\nabla\langle \boldsymbol{u} \rangle)^T}{2} = \frac{\rho\nu}{2} \begin{pmatrix} 0 & \partial_y \langle u_1 \rangle & 0 \\ \partial_y \langle u_1 \rangle & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The mean viscous stress at the wall, the so-called wall shear stress, is the non-vanishing component

$$\tau_{\mathrm{w}} = \rho\nu \partial_y \langle u_1 \rangle \big|_{y=0} \quad [\mathrm{N}/\mathrm{m}^2],$$

where the unit assumes a velocity with dimension. With τ_{w}, one can define an appropriate velocity and length scale for the near-wall region, which enable a characterization of the layer independently of the Reynolds number. The velocity scale is the so-called friction velocity given by

$$U_\tau = \sqrt{\frac{\tau_{\mathrm{w}}}{\rho}} \quad [\mathrm{m}/\mathrm{s}]$$

and the viscous length scale is

$$\lambda_\nu = \nu \sqrt{\frac{\rho}{\tau_{\mathrm{w}}}} = \frac{\nu}{U_\tau} \quad [\mathrm{m}].$$

Then, the non-dimensional distance from the wall can be measured in viscous length or wall units

$$y^+ = \frac{y}{\lambda_\nu} = \frac{U_\tau}{\nu} y.$$

Depending on y^+, several regions for near-wall flows are distinguished. The region $y^+ < 50$ is called viscous wall region since effects of the molecular viscosity are of importance. For $y^+ < 5$, the viscous stress τ_{w} even dictates the behavior of the flow in this region and it is called viscous sublayer. The region $y^+ > 50$ is called outer layer. □

Remark 8.14 (Differences Between Two- and Three-dimensional Flows) There are at least two fundamental differences between two- and three-dimensional flows.

First, the smallest scales in two-dimensional flows behave differently than (8.7). In Kraichnan (1967) it was shown that they are of order $\lambda_{2d} \sim \mathrm{Re}^{-1/2}$ in two dimensions, where λ_{2d} is the Kraichnan dissipation length.

A second difference is the so-called vortex stretching. The vorticity is defined by $\boldsymbol{\omega} = \nabla \times \boldsymbol{u}$, see Definition 3.162. Applying the curl operator to the Navier–Stokes equations (8.1) and assuming sufficiently smooth functions gives for the viscous term, applying (2.27), the Theorem of Schwarz, and again (2.27), and using $\nabla \cdot \boldsymbol{\omega} = 0$ which follows from (3.157),

$$\begin{aligned}
2\nabla \times (\nabla \cdot \mathbb{D}(\boldsymbol{u})) = \nabla \times (\Delta \boldsymbol{u}) &= \begin{pmatrix} \partial_y \Delta u_3 - \partial_z \Delta u_2 \\ \partial_z \Delta u_1 - \partial_x \Delta u_3 \\ \partial_x \Delta u_2 - \partial_y \Delta u_1 \end{pmatrix} \\
&= \begin{pmatrix} \Delta\left(\partial_y u_3 - \partial_z u_2\right) \\ \Delta\left(\partial_z u_1 - \partial_x u_3\right) \\ \Delta\left(\partial_x u_2 - \partial_y u_1\right) \end{pmatrix} = \Delta(\nabla \times \boldsymbol{u}) \\
&= \Delta \boldsymbol{\omega} = 2\nabla \cdot \mathbb{D}(\boldsymbol{w}).
\end{aligned}$$

For the convective term, one obtains with (6.6), (3.156), and (3.159), using again that $\nabla \cdot \boldsymbol{\omega} = 0$,

$$\begin{aligned}
\nabla \times ((\boldsymbol{u} \cdot \nabla)\boldsymbol{u}) &= \nabla \times \left((\nabla \times \boldsymbol{u}) \times \boldsymbol{u} + \frac{1}{2}\nabla\left(\boldsymbol{u}^T\boldsymbol{u}\right)\right) = \nabla \times (\boldsymbol{\omega} \times \boldsymbol{u}) \\
&= (\boldsymbol{u} \cdot \nabla)\boldsymbol{\omega} - (\boldsymbol{\omega} \cdot \nabla)\boldsymbol{u}.
\end{aligned}$$

Thus, one gets for the Navier–Stokes equations with $\boldsymbol{f} = \boldsymbol{0}$ the following equation for the vorticity

$$\frac{\partial \boldsymbol{\omega}}{\partial t} - 2\mathrm{Re}^{-1}\nabla \cdot \mathbb{D}(\boldsymbol{\omega}) + (\boldsymbol{u} \cdot \nabla)\boldsymbol{\omega} - (\boldsymbol{\omega} \cdot \nabla)\boldsymbol{u} = \boldsymbol{0}.$$

The viscous term is small for high Reynolds numbers and can be neglected. Thus

$$\frac{D\boldsymbol{\omega}}{Dt} = \frac{\partial \boldsymbol{\omega}}{\partial t} + (\boldsymbol{u} \cdot \nabla)\boldsymbol{\omega} \approx (\boldsymbol{\omega} \cdot \nabla)\boldsymbol{u}. \tag{8.8}$$

This relation is the equation of an infinitesimal line element of material. If $\nabla \boldsymbol{u}$ acts to stretch the line element, then $|\boldsymbol{\omega}|$ will be stretched, too. In fact, in turbulent three-dimensional flows, vortex stretching occurs, which is an important feature of such flows. Vortex stretching causes a change of the local length scale and it leads to a hierarchy of vortical structures of different size. Thus, vortex stretching is responsible for the multiscale character of turbulent flows.

In two-dimensional flows, the right-hand side of (8.8) vanishes, which induces that vortex stretching cannot occur.

Because of the absence of vortex stretching and the different size of the viscous length scales, two-dimensional flows at high Reynolds numbers are qualitatively different to three-dimensional turbulent flows. For this reason, one can share the point of view that in two dimensions there are no turbulent flows. However, it seems to be legitimate to check new methods for high Reynolds number flows also at two-dimensional problems. If they fail, their success at three-dimensional flows, which possess additional complex features, is unlikely. On the other hand, if they are successful, it cannot be concluded without numerical studies that they will work well in three dimensions, too. □

Remark 8.15 (A Mathematical Approach for Studying Turbulent Flows) A mathematical concept for studying turbulent flows was developed within the theory of dynamical systems. A dynamical system is given by

$$\frac{d\boldsymbol{u}}{dt} = F(\boldsymbol{u}), \quad \boldsymbol{u}_0 = \boldsymbol{u}(0). \tag{8.9}$$

The incompressible Navier–Stokes equations in the weakly divergence-free space (7.3) fit into this concept with

$$F(\boldsymbol{u}) = \boldsymbol{f} - \nu A\boldsymbol{u} - N(\boldsymbol{u}, \boldsymbol{u}),$$

where $A : V_{\mathrm{div}} \to V'$ is the Laplace operator in the distributional sense defined by

$$(\nabla \boldsymbol{u}, \nabla \boldsymbol{v}) = \langle A\boldsymbol{u}, \boldsymbol{v}\rangle_{V',V} \quad \forall\ \boldsymbol{v} \in V_{\mathrm{div}}$$

and $N : V_{\mathrm{div}} \times V_{\mathrm{div}} \to V'$ is the bilinear convective operator

$$\left(N\left(\boldsymbol{u}, \boldsymbol{v}\right), \boldsymbol{w}\right) = n\left(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\right) \quad \forall\ \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V_{\mathrm{div}}.$$

In the theory for the Navier–Stokes equations it is assumed that the body force is independent of time, i.e., $\boldsymbol{f}(t, \boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})$. In this case, one obtains an autonomous dynamical system.

To study sets of solutions of dynamical systems, one can define attractors, see Remark 8.18, and in particular the dimension of these attractors. In Temam (1997, p. 380) it is written 'It is our understanding here that the number of degrees of freedom of a turbulent phenomenon is the dimension of the attractor which represents it'.

Here, a short introduction into this concept and a brief review of some results for the incompressible Navier–Stokes equations will be given. More detailed presentations can be found, e.g., in Foias et al. (2001, Chap. III) or Marion and Temam (1998, Chap. 13). □

Remark 8.16 (A Family of Operators Describing the Evolution of a Dynamical System) Consider the autonomous dynamical system (8.9) that is assumed to possess a unique solution. Then, a family of operators $\{S(t)\}_{t\geq 0}$ is defined by

$$S(t) \,:\, \boldsymbol{u}(0) \mapsto \boldsymbol{u}(t) \qquad \Longleftrightarrow \qquad \boldsymbol{u}(t) = S(t)\boldsymbol{u}(0),$$

i.e., the initial condition is mapped to the solution at time t. Basic properties are the followings:

- $S(0) = \mathrm{I}$, where I is the identity operator.
- One obtains the same solution at time $t+s$, $t, s \geq 0$,
  - if the system is started at time 0 and evolved until time $t+s$ or
  - if the system is started at time 0, evolved until time t and then evolved further until time $t+s$ is reached.

  Mathematically, this property is

  $$\boldsymbol{u}(t+s) = S(t+s)\boldsymbol{u}(0) = S(s)\left(S(t)\boldsymbol{u}(0)\right) = \left(S(s)\circ S(t)\right)\boldsymbol{u}(0) \quad \forall\ s,t \geq 0$$

  or

  $$S(t+s) = S(s)\circ S(t) = S(t)\circ S(s) \quad \forall\ s,t \geq 0.$$

Therefore, the operators $\{S(t)\}_{t\geq 0}$ form a semigroup. There are no inverse elements. □

Remark 8.17 (Setup for the Incompressible Navier–Stokes Equations) For the incompressible Navier–Stokes equations, it is assumed that $\boldsymbol{u}_0, \boldsymbol{f} \in H_{\mathrm{div}}(\Omega)$. Note that the assumption of unique solvability for the incompressible Navier–Stokes equations is known so far only for two dimensions, see Theorem 7.24. In this case, one can show that $S(t)$ is a continuous operator in $H_{\mathrm{div}}(\Omega)$, Marion and Temam (1998, p. 554), Foias et al. (2001, p. 138). □

Remark 8.18 (Global Attractor, Foias et al. (2001, p. 138)) The global attractor of $\{S(t)\}_{t\geq 0}$ is a set $\mathcal{A} \subset H_{\mathrm{div}}(\Omega)$ with the following properties:

- The set $\mathcal{A}$ is compact in $H_{\mathrm{div}}(\Omega)$.
- The set $\mathcal{A}$ is invariant for the semigroup, i.e., $S(t)\mathcal{A} = \mathcal{A}$ for all $t \geq 0$.
- The set $\mathcal{A}$ attracts all bounded sets in $H_{\mathrm{div}}(\Omega)$, i.e., for every bounded set $B \subset H_{\mathrm{div}}(\Omega)$ it is

  $$\mathrm{dist}_{H_{\mathrm{div}}(\Omega)}\left(S(t)B, \mathcal{A}\right) = \sup_{b\in S(t)B}\inf_{a\in\mathcal{A}} \|a-b\|_{L^2(\Omega)} \to 0 \quad \text{as} \quad t\to\infty,$$

  i.e., the distant between these two sets goes to zero for $t \to \infty$.

If the global attractor exists, it can be shown that it is unique. □

Remark 8.19 (The Hausdorff Dimension of the Global Attractor, Foias et al. (2001, p. 117)) The so-called Hausdorff dimension of a global attractor is of particular interest for the incompressible Navier–Stokes equations. Given a set X and $\varepsilon > 0$. This set is covered by balls of dimension $\hat{d} \in \mathbb{R}$ with radii that are not larger than ε. The volume of these balls is proportional to $\varepsilon^{\hat{d}}$. Defining the quantity

$$\mu_{\mathrm{H}}(X, \hat{d}, \varepsilon) = \inf \sum_{i \in I} r_i^{\hat{d}},$$

where the infimum is for all coverings of X with a family $\{B_i\}_{i \in I}$ of balls with radius $r_i \le \varepsilon$. Reducing the maximal radius, the volume of the covering will not increase, thus $\mu_{\mathrm{H}}(X, \hat{d}, \varepsilon)$ is a non-increasing function with respect to ε. Then,

$$\mu_{\mathrm{H}}(X, \hat{d}) = \lim_{\varepsilon \to 0} \mu_{\mathrm{H}}(X, \hat{d}, \varepsilon)$$

is the $\hat{d}$-dimensional Hausdorff measure of X. One can show that there is a number $d_0 \in [0, \infty]$ such that $\mu_{\mathrm{H}}(X, \hat{d}) = \infty$ for $\hat{d} < d_0$ and $\mu_{\mathrm{H}}(X, \hat{d}) = 0$ for $\hat{d} > d_0$. This number d_0 is called the Hausdorff dimension of X and it is denoted by $d_{\mathrm{H}}(X)$. □

Remark 8.20 (Results for the Two-dimensional Navier–Stokes Equations) Since there exists a unique solution of the incompressible Navier–Stokes equations in two dimensions, one can define the semigroup $\{S(t)\}_{t \ge 0}$ and study its properties. In Foiaş and Temam (1979), the existence of a global attractor and the finiteness of its Hausdorff dimension were proved in the case of a smooth boundary of Ω. There are several extensions of these results, in particular to the periodic case, see Foias et al. (2001, p. 140) for an overview. An estimate of the Hausdorff dimension was given in Temam (1986). It can be shown that

$$d_{\mathrm{H}}(\mathcal{A}) \sim \left(\frac{L}{\lambda_{\mathrm{2d}}} \right)^2,$$

see Foias et al. (2001, p. 141), where λ_{2d} is introduced in Remark 8.14. □

Remark 8.21 (Extensions to the Navier–Stokes Equations in Three Dimensions) For the three-dimensional Navier–Stokes equations, one can consider so-called invariant sets which are bounded in V_{div}, see Foias et al. (2001, p. 147). A set X in V_{div} is invariant if, for any initial condition $\boldsymbol{u}_0 \in X$, the corresponding unique local solution, see Remark 7.26, extends globally in time to a unique solution $\boldsymbol{u}(t)$ that is defined for all $t \ge 0$ with values in X. It was shown in Constantin et al. (1985b) and Constantin et al. (1985a) that the Hausdorff dimension of any invariant bounded set $X \subset V_{\mathrm{div}}$ has the same order concerning the number of degrees of freedom as predicted from Kolmogorov's theory, see Remark 8.11, i.e.,

$$d_{\mathrm{H}}(X) \sim \left(\frac{L}{\lambda} \right)^3. \tag{8.10}$$

□

Remark 8.22 (Smaller Complexity of Turbulence Models) As already mentioned in Remark 8.1, a turbulence model has to be less complex than the incompressible Navier–Stokes equations. Each turbulence model has a parameter that determines the scales that should be simulated. This parameter is usually related to the (local) mesh width, e.g., as in (8.67) for the Smagorinsky LES model, since the mesh width determines which scales can be represented. Thus, in practice, the parameter becomes smaller on finer meshes and asymptotically the turbulence model converges to the Navier–Stokes equations as the mesh width tends to zero. To study if a turbulence model is less complex than the Navier–Stokes equations, one fixes the parameter, like in Sect. 8.3.2, and considers the continuous counterpart of the turbulence model. Indicators for a reduced complexity are the existence and uniqueness of a weak solution for long times since the uniqueness is not proved for the Navier–Stokes equations so far, see Remark 7.25, or that the dimension of the global attractor is smaller than (8.10).

Another indicator, coming from finite element error analysis, is that one can prove error estimates which depend not on the Reynolds number but on some reduced Reynolds number, e.g., see Theorem 8.273 for a three-scale coarse space projection-based variational multiscale method. □

8.2 Large Eddy Simulation: The Concept of Space Averaging

8.2.1 The Basic Concept of LES, Space Averaging, Convolution with Filters

Remark 8.23 (The Basic Idea of Large Eddy Simulation) In the case of turbulent flows, only, in some sense, large scales can be represented on given grids and these scales are the only ones that can be simulated, see Remark 8.11 for a discussion of this issue. There are different concepts of defining scales to be large. In the case of Large Eddy Simulation (LES), the large scales are defined by an average in space. In theory, the space averaging is usually given by a convolution with a filter function.

With this approach, one obtains a decomposition

$$\boldsymbol{u} = \overline{\boldsymbol{u}} + \boldsymbol{u}', \quad p = \overline{p} + p', \tag{8.11}$$

of the solution of the Navier–Stokes equations. In (8.11), $(\overline{\boldsymbol{u}}, \overline{p})$ are the large scales, and $(\boldsymbol{u}', p')$ are the small scales or unresolved scales or subgrid scales or fluctuations. □

Remark 8.24 (Space Averaging with Convolution, Filter Function, Wave Numbers) The space average of a sufficiently smooth function $v(t, \boldsymbol{x})$ defined in $\mathbb{R}^d$ given by a convolution with an appropriate filter function $g_{\mathrm{fil}}(\boldsymbol{x})$ is defined by

$$\begin{aligned} \overline{v}\ (t, \boldsymbol{x}) &= (g_{\mathrm{fil}} * v)\ (t, \boldsymbol{x}) = \int_{\mathbb{R}^d} g_{\mathrm{fil}}\ (\boldsymbol{x} - \boldsymbol{z})\ v\ (t, \boldsymbol{z})\ d\boldsymbol{z} \\ &= \int_{\mathbb{R}^d} g_{\mathrm{fil}}\ (\boldsymbol{z})\ v\ (t, \boldsymbol{x} - \boldsymbol{z})\ d\boldsymbol{z}. \end{aligned} \tag{8.12}$$

This identity follows from a transform of the integration variable. Consider for simplicity the one-dimensional case, one sets first $w = x - z$ and then one switches the bounds for the integration to obtain

$$\int_{-\infty}^{\infty} g_{\mathrm{fil}}(x - z)v(z)\ dz = -\int_{\infty}^{-\infty} g_{\mathrm{fil}}(w)v(x - w)\ dw = \int_{-\infty}^{\infty} g_{\mathrm{fil}}(z)v(x - z)\ dz.$$

Applying the Fourier transform to (8.12) gives with (A.25)

$$\mathcal{F}\left(\overline{v}\right)(t, \boldsymbol{k}) = \left(\mathcal{F}\left(g_{\mathrm{fil}}\right)\mathcal{F}\left(v\right)\right)(t, \boldsymbol{k}), \tag{8.13}$$

where $\boldsymbol{k}$ is the dual variable or wave number. If $\mathcal{F}\left(g_{\mathrm{fil}}\right)(t, \boldsymbol{k}) = 0$ for $\|\boldsymbol{k}\|_2 > k_c$, where k_c is a cut-off wave number, then all high wave number components of $v\ (t, \boldsymbol{x})$ are filtered out by convolving v with g_{fil}. The high wave number components correspond to the fluctuations of v. This situation is the ideal one. For practical applications, it is sufficient that the high wave numbers are damped sufficiently fast. Thus, an essential requirement on a filter function is that its Fourier transform decreases rapidly for high wave numbers.

It will be assumed that the filter function $g_{\mathrm{fil}}\ (\boldsymbol{x})$ can be represented as tensor product of one-dimensional filter functions

$$g_{\mathrm{fil}}\ (\boldsymbol{x}) = \prod_{i=1}^{d} g_{\mathrm{fil},i}\ (x_i)\ . \tag{8.14}$$

Since all factors have different variables, the Fourier transform of the filter function is

$$\mathcal{F}\left(g_{\mathrm{fil}}\right)(\boldsymbol{k}) = \prod_{i=1}^{d} \mathcal{F}\left(g_{\mathrm{fil},i}\right)(k_i)\ .$$

The filtering effect in the definition of $\overline{v}$ is often described by a positive constant δ, called the characteristic filter width or averaging radius or scale of the filter. The filter width is a measure of the size of the eddies that are damped out, which are all eddies with size less than $\mathcal{O}\left(\delta\right)$. It is clear that the smaller δ is the larger becomes the cut-off wave number k_c and the less eddies are filtered out. □

Example 8.25 (The Gaussian Filter) The Gaussian filter is one of the most common filters applied in LES. Its filter function is given by

$$g_{\mathrm{fil},i}(x_i) = g_{\mathrm{Gauss},i}(x_i) = \sqrt{\frac{6}{\pi\delta^2}} \exp\left(-\frac{6}{\delta^2}x_i^2\right). \tag{8.15}$$

It is

$$\int_0^\infty \exp(-ax^2)\,dx = \frac{1}{2}\sqrt{\frac{\pi}{a}}, \quad a > 0,$$

and because of symmetry

$$\int_{-\infty}^\infty \exp(-ax^2)\,dx = \sqrt{\frac{\pi}{a}}, \quad a > 0. \tag{8.16}$$

Then, one obtains for the Fourier transform (A.23)

$$\begin{aligned}
&\mathcal{F}\left(g_{\mathrm{Gauss},j}\right)\left(k_j\right) \\
&= \sqrt{\frac{6}{\pi\delta^2}} \int_{-\infty}^\infty \exp\left(-\frac{6}{\delta^2}x_j^2 - ik_jx_j\right)\,dx_j \\
&= \sqrt{\frac{6}{\pi\delta^2}} \int_{-\infty}^\infty \exp\left(-\left(\frac{\sqrt{6}}{\delta}x_j + \frac{ik_j\delta}{2\sqrt{6}}\right)^2 + \frac{i^2k^2\delta^2}{24}\right)\,dx_j \\
&= \sqrt{\frac{6}{\pi\delta^2}} \exp\left(-\frac{k_j^2\delta^2}{24}\right) \int_{-\infty}^\infty \exp\left(-\left(\frac{\sqrt{6}}{\delta}x_j + \frac{ik_j\delta}{2\sqrt{6}}\right)^2\right)\,dx_j \\
&= \sqrt{\frac{6}{\pi\delta^2}} \frac{\delta}{\sqrt{6}} \exp\left(-\frac{\delta^2}{24}k_j^2\right) \int_{-\infty}^\infty \exp\left(-y^2\right)\,dy \\
&= \exp\left(-\frac{\delta^2}{24}k_j^2\right), \quad j = 1, \ldots, d.
\end{aligned} \tag{8.17}$$

The factor in (8.15) is chosen such that the filtering effect of the Gaussian filter and the filtering effect of a certain discrete approximation in a model problem are equilibrated, see Aldama (1990, Sect. 3.8) for a detailed discussion. In addition, the integral of the filter function is normalized, see (8.21). The Gaussian filter function and its Fourier transform are of the same form.

The Gaussian filter g_{Gauss} in $\mathbb{R}^d$ is given by

$$g_{\mathrm{Gauss}}(\boldsymbol{x}) = \left(\frac{6}{\delta^2\pi}\right)^{d/2} \exp\left(-\frac{6}{\delta^2}\|\boldsymbol{x}\|_2^2\right), \tag{8.18}$$

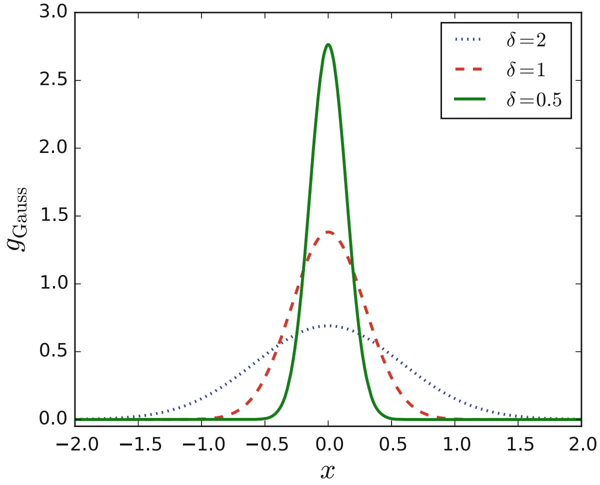

Fig. 8.2 The Gaussian filter in one dimension for different values of the filter width δ

see Fig. 8.2. Its Fourier transform is

$$\mathcal{F}\left(g_{\text{Gauss}}\right)(\boldsymbol{k}) = \exp\left(-\frac{\delta^2}{24}\|\boldsymbol{k}\|_2^2\right). \tag{8.19}$$

For convenience of notation, the Gaussian filter with a scalar argument x is understood to be

$$g_{\text{Gauss}}(x) := \left(\frac{6}{\delta^2\pi}\right)^{\frac{d}{2}} \exp\left(-\frac{6x^2}{\delta^2}\right).$$

The Gaussian filter has the following properties, which follow directly from properties of the exponential or which can be verified by direct calculations, compare also Fig. 8.2:

- *regularity:* Since the exponential is infinitely often differentiable, it follows that

$$g_{\text{Gauss}} \in C^\infty\left(\mathbb{R}^d\right), \quad \mathcal{F}\left(g_{\text{Gauss}}\right) \in C^\infty\left(\mathbb{R}^d\right).$$

- *positivity:* The positivity of the exponential and the fact that (8.18) and (8.19) take their maximal value at the origin yields

$$0 < g_{\text{Gauss}}(\boldsymbol{x}) \le \left(\frac{6}{\delta^2\pi}\right)^{\frac{d}{2}}, \quad 0 < \mathcal{F}\left(g_{\text{Gauss}}\right)(\boldsymbol{k}) \le 1. \tag{8.20}$$

- *integrability:* From (8.16), one finds that

$$\|g_{\mathrm{Gauss}}\|_{L^p(\mathbb{R}^d)} = \left(\frac{6}{\delta^2\pi}\right)^{d/2} \left(\int_{\mathbb{R}^d} \exp\left(-p\frac{6}{\delta^2}\|\boldsymbol{x}\|_2^2\right)\,dx\right)^{1/p}$$
$$= \left(\frac{6}{\delta^2\pi}\right)^{d/2} \left(\frac{\pi\delta^2}{6p}\right)^{d/(2p)}.$$

Together with (8.20), one gets

$$\|g_{\mathrm{Gauss}}\|_{L^p(\mathbb{R}^d)} < \infty,\ p \in [1,\infty],, \quad \|g_{\mathrm{Gauss}}\|_{L^1(\mathbb{R}^d)} = 1. \tag{8.21}$$

- *symmetry:* Since the functions (8.18) and (8.19) depend only on the Euclidean norm of $\boldsymbol{x}$ and $\boldsymbol{k}$, it follows that

$$g_{\mathrm{Gauss}}(\boldsymbol{x}) = g_{\mathrm{Gauss}}(-\boldsymbol{x}), \quad \mathcal{F}(g_{\mathrm{Gauss}})(\boldsymbol{k}) = \mathcal{F}(g_{\mathrm{Gauss}})(-\boldsymbol{k}). \tag{8.22}$$

- *monotonicity:* From the monotonicity of the exponential, one obtains

$$g_{\mathrm{Gauss}}(\boldsymbol{x}) \ge g_{\mathrm{Gauss}}(\boldsymbol{y}) \quad \text{if} \quad \|\boldsymbol{x}\|_2 \le \|\boldsymbol{y}\|_2. \tag{8.23}$$

□

Lemma 8.26 (Further Properties of the Gaussian Filter)

i) Let $\varphi \in L^p(\mathbb{R}^d)$, then for $1 \le p < \infty$

$$\lim_{\delta\to 0}\|g_{\mathrm{Gauss}} * \varphi - \varphi\|_{L^p(\mathbb{R}^d)} = 0. \tag{8.24}$$

ii) Let $\varphi \in L^\infty(\mathbb{R}^d)$. If φ is uniformly continuous on a set ω, then there is a pointwise convergence $g_{\mathrm{Gauss}} * \varphi \to \varphi$ uniformly on ω as $\delta \to 0$.

Proof The proof of the first two statements can be found, e.g., in Folland (1995, Theorem 0.13). ■

Example 8.27 (The Box Filter) Another popular filter is the box filter or top hat filter. It is given in one dimension for the filter width $\delta > 0$ by

$$g_{\mathrm{box}} = \begin{cases} \frac{1}{\delta} & \text{for } x \in \left[-\frac{\delta}{2}, \frac{\delta}{2}\right], \\ 0 & \text{else,} \end{cases}$$

see Fig. 8.3. For multiple dimensions, it is defined by a tensor product as in (8.14). Main properties of the box filter are as follows:

- it is piecewise constant,
- it is non-negative,
- it has compact support: $\mathrm{supp}(g_{\mathrm{box}}) = [-\delta/2, \delta/2]$,

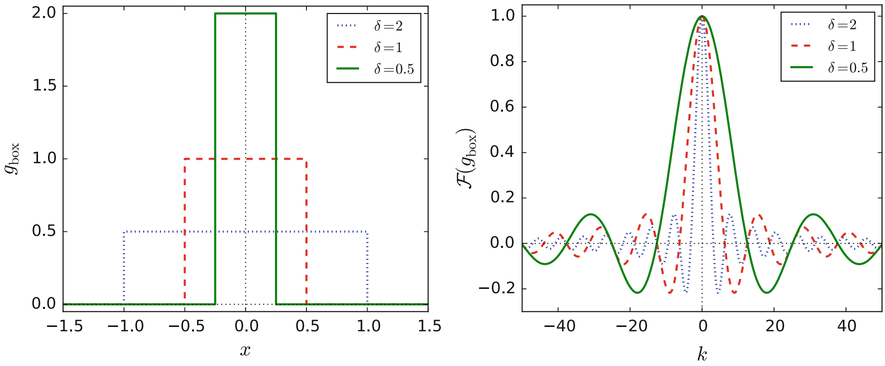

Fig. 8.3 The box filter (*left*) and its Fourier transform (*right*) in one dimension for different values of the filter width δ

- $g_{\text{box}} \in L^p(\mathbb{R})$ for $p \in [1, \infty]$, $\|g_{\text{box}}\|_{L^1(\mathbb{R})} = 1$.
- For the Fourier transform (A.23), one obtains with the fundamental theorem of calculus and with Euler's formula for the exponential with complex argument

$$
\begin{aligned}
&\mathcal{F}\left(g_{\text{box},j}\right)\left(k_j\right) \\
&= \int_{-\delta/2}^{\delta/2} \frac{1}{\delta} \exp\left(-ik_j x_j\right)\, dx_j \\
&= \frac{1}{\delta}\left(-\frac{1}{ik_j}\right)\left(\exp\left(-\frac{ik_j\delta}{2}\right) - \exp\left(\frac{ik_j\delta}{2}\right)\right) \\
&= -\frac{1}{ik_j\delta}\left(\cos\left(\frac{k_j\delta}{2}\right) - i\sin\left(\frac{k_j\delta}{2}\right) - \cos\left(\frac{k_j\delta}{2}\right) - i\sin\left(\frac{k_j\delta}{2}\right)\right) \\
&= \frac{2}{k_j\delta}\sin\left(\frac{k_j\delta}{2}\right).
\end{aligned}
$$

Thus, the Fourier transform of the box filter is a damped sine function. This function has negative values, see Fig. 8.3.

□

8.2.2 The Space-Averaged Navier–Stokes Equations in the Case $\Omega = \mathbb{R}^d$

Remark 8.28 (Steps for Deriving LES Models) To compute the space-averaged velocity $\overline{\boldsymbol{u}}$ and pressure $\overline{p}$, equations for these quantities are needed. These equations have to be derived from the governing equations for $\boldsymbol{u}$ and p, i.e., from the Navier–Stokes equations. The first step consists in applying the filter which defines $(\overline{\boldsymbol{u}}, \overline{p})$ also to the Navier–Stokes equations (2.25). Thus, LES modeling is based on the strong form of the Navier–Stokes equations, which is in contrast to other turbulence models, which are based on the variational form of the Navier–Stokes equations, e.g., see Sect. 8.8. After having applied the filter to (2.25), the basic equation for LES, the space-averaged Navier–Stokes equations, are obtained with the assumption that convolution and differentiation commute. However, due to the nonlinear term of the Navier–Stokes equations, an additional modeling step is necessary to derive equations for $(\overline{\boldsymbol{u}}, \overline{p})$ from the space-averaged Navier–Stokes equations.

The assumption that convolution and differentiation commute will be studied in the remainder of Sect. 8.2. The additional modeling step is the topic of Sects. 8.3 and 8.4. □

Remark 8.29 (Assumptions on the Filter) It will be assumed that the filter operator which defines $(\overline{\boldsymbol{u}}, \overline{p})$ has the following properties:

- The filter is a linear operator

$$\overline{\boldsymbol{u} + \lambda \boldsymbol{v}} = \overline{\boldsymbol{u}} + \lambda\, \overline{\boldsymbol{v}}. \tag{8.25}$$

- Derivatives (first and second order in space, first order in time) and averages commute, e.g.,

$$\overline{\partial_{x_i} \boldsymbol{u}} = \partial_{x_i} \overline{\boldsymbol{u}}, \quad i = 1, \ldots, d, \tag{8.26}$$

and

$$\overline{\partial_t \boldsymbol{u}} = \partial_t \overline{\boldsymbol{u}}. \tag{8.27}$$

It is known already from classical calculus that the commutation of differentiation and integration of a parametric function requires sufficient smoothness of this function with respect to the parameter.

□

Remark 8.30 (Filtering with Convolution in the Case $\Omega = \mathbb{R}^d$) Let $\Omega = \mathbb{R}^d$ and consider as filter operator the convolution with a filter function given in (8.12). Note that this filter operator is only defined in $\mathbb{R}^d$. The linearity (8.25) follows from the linearity of the integral operator. The commutation with respect to the temporal derivative (8.27) follows from the fact that integration with respect to a certain variable and differentiation with respect to another variable can be interchanged. For property (8.26), it will be assumed in addition that the filter width δ is constant. Then, one gets for sufficiently smooth functions with (8.12), the transform of the integration variable $\boldsymbol{w} = \boldsymbol{x} - \boldsymbol{z}$, Fubini's theorem, the back transform of integration variable, and the chain rule to compute

$$\partial_{x_i} v = \sum_{j=1}^{d} \partial_{z_j} v \partial_{x_i} z_j = \sum_{j=1}^{d} \partial_{z_j} v \delta_{ij} = \partial_{z_i} v,$$

for $i = 1, \ldots, d$,

$$\begin{aligned}
\partial_{x_i} \int_{\mathbb{R}^d} g_{\text{fil}} (\boldsymbol{x} - \boldsymbol{z}) \, v(t, \boldsymbol{z}) \, d\boldsymbol{z} &= -\partial_{x_i} \int_{\infty}^{-\infty} \cdots \int_{\infty}^{-\infty} g_{\text{fil}} (\boldsymbol{w}) \, v(t, \boldsymbol{x} - \boldsymbol{w}) \, d\boldsymbol{w} \\
&= -\int_{\infty}^{-\infty} \cdots \int_{\infty}^{-\infty} g_{\text{fil}} (\boldsymbol{w}) \, \partial_{x_i} v(t, \boldsymbol{x} - \boldsymbol{w}) \, d\boldsymbol{w} \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g_{\text{fil}} (\boldsymbol{x} - \boldsymbol{z}) \, \partial_{x_i} v(t, \boldsymbol{z}) \, d\boldsymbol{z} \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g_{\text{fil}} (\boldsymbol{x} - \boldsymbol{z}) \, \partial_{z_i} v(t, \boldsymbol{z}) \, d\boldsymbol{z}.
\end{aligned}$$

□

Remark 8.31 (The Space-Averaged Navier–Stokes Equations in the Case $\Omega = \mathbb{R}^d$) Applying a filter with the properties (8.25) and (8.26) to the Navier–Stokes equations (2.25) with sufficiently smooth functions and the initial condition $\boldsymbol{u}_0$ gives the space-averaged Navier–Stokes equations (or Reynolds equations)

$$\begin{aligned}
\partial_t \overline{\boldsymbol{u}} - 2\nu \nabla \cdot \mathbb{D}(\overline{\boldsymbol{u}}) + \nabla \cdot \left(\overline{\boldsymbol{u}\boldsymbol{u}^T} \right) + \nabla \overline{p} = \overline{\boldsymbol{f}} \quad & \text{in } (0, T] \times \mathbb{R}^d, \\
\nabla \cdot \overline{\boldsymbol{u}} = 0 \quad & \text{in } (0, T] \times \mathbb{R}^d, \\
\overline{\boldsymbol{u}}(0, \cdot) = \overline{\boldsymbol{u}}_0 \quad & \text{in } \mathbb{R}^d.
\end{aligned} \tag{8.28}$$

□

Remark 8.32 (The Filter of the Nonlinear Term and the Closure Problem) The space-averaged Navier–Stokes equations (8.28) are not yet equations for $(\overline{\boldsymbol{u}}, \overline{p})$ because the tensor $\overline{\boldsymbol{u}\boldsymbol{u}^T}$ is not expressed in terms of $(\overline{\boldsymbol{u}}, \overline{p})$. Since the dyadic product of a d-dimensional vector with itself is a symmetric tensor, $\overline{\boldsymbol{u}\boldsymbol{u}^T}$ is symmetric, too. Thus, on the one hand there are $(d + 1)$ equations in (8.28) and on the other hand there are $(d + 1)$ unknown space-averaged values and $d(d + 1)/2$ unknown quantities in $\overline{\boldsymbol{u}\boldsymbol{u}^T}$. Hence, a closure problem arises. Writing the nonlinear term in the form

$$\nabla \cdot \left(\overline{\boldsymbol{u}\boldsymbol{u}^T} \right) = \nabla \cdot \left(\overline{\boldsymbol{u}}\, \overline{\boldsymbol{u}}^T \right) + \nabla \cdot \left(\overline{\boldsymbol{u}\boldsymbol{u}^T} - \overline{\boldsymbol{u}}\, \overline{\boldsymbol{u}}^T \right) = \nabla \cdot \left(\overline{\boldsymbol{u}}\, \overline{\boldsymbol{u}}^T \right) + \nabla \cdot \mathbb{T}, \tag{8.29}$$

where $\mathbb{T}$ is called subgrid-scale (sgs) stress tensor, Meneveau and Katz (2000), subgrid tensor, Sagaut (2006, p. 49), or residual-stress tensor, Pope (2000, p. 581), or Reynolds stress tensor. A model of the subgrid scale tensor is needed for closing the equations. Modeling $\mathbb{T}$ in terms of $(\overline{\boldsymbol{u}}, \overline{p})$ is the main issue in LES. This topic will be addressed in Sects. 8.3 and 8.4. □

Remark 8.33 (Averaging in Time) The concept of LES models does not include averaging in time. Temporal averaging is used in the derivation of Reynolds Averaged Navier–Stokes (RANS) models, e.g., see Ferziger and Perić (1999, Sect. 9.4) for more details. □

8.2.3 The Space-Averaged Navier–Stokes Equations in a Bounded Domain

Remark 8.34 (The Space-Averaged Navier–Stokes Equations in Bounded Domains in Practical Computations) Usually, (8.28) is also used in practical computations in bounded domains Ω, i.e., $\mathbb{R}^d$ is simply replaced by Ω, also in the definition (8.12) of the filter. However, if Ω is a bounded domain, the commutation of filtering and differentiation requires special attention from the mathematical point of view. One has to extend all functions of the Navier–Stokes equations consistently from Ω to $\mathbb{R}^d$. In general, these extensions are not sufficiently smooth for the simple commutation of differentiation and integration to be valid. An extra term occurs and omitting this term leads to a so-called commutation error. This type of commutation error will be studied in the following.

The presentation of this topic is based on Dunca et al. (2004). □

Remark 8.35 (Setup of the Problem) Let Ω be a bounded domain in $\mathbb{R}^d$, $d \in \{2, 3\}$, with Lipschitz boundary $\partial\Omega$ and $(d-1)$-dimensional measure $|\partial\Omega| < \infty$. The incompressible Navier–Stokes equations with homogeneous Dirichlet boundary conditions

$$\begin{aligned} \partial_t \boldsymbol{u} - 2\nu\nabla \cdot \mathbb{D}(\boldsymbol{u}) + \nabla \cdot \left(\boldsymbol{u}\boldsymbol{u}^T\right) + \nabla p &= \boldsymbol{f} \quad \text{in } (0,T] \times \Omega, \\ \nabla \cdot \boldsymbol{u} &= 0 \quad \text{in } (0,T] \times \Omega, \\ \boldsymbol{u} &= \boldsymbol{0} \quad \text{on } (0,T] \times \partial\Omega, \\ \boldsymbol{u}(0,\cdot) &= \boldsymbol{u}_0 \text{ in } \Omega, \\ \int_\Omega p \, d\boldsymbol{x} &= 0 \quad \text{in } (0,T], \end{aligned} \tag{8.30}$$

will be considered.

The derivation of the space-averaged Navier–Stokes equations and the analysis of the commutation error require that (8.30) possesses a unique solution $(\boldsymbol{u}, p)$ which is sufficiently regular, such that the normal stress has a well defined trace on $\partial\Omega$ that belongs to some Lebesgue space defined on $\partial\Omega$. The stress tensor $\mathbb{S}$ is given in (2.35) and the normal stress or Cauchy stress vector on $\partial\Omega$ is defined by $\mathbb{S}\boldsymbol{n}$. Concretely, it will be assumed that

$$\begin{aligned} \boldsymbol{u} \in H^2(\Omega) \cap V,\ p \in H^1(\Omega) \cap Q \text{ for a.e. } t \in [0,T], \\ \boldsymbol{u} \in H^1((0,T)) \text{ for a.e. } \boldsymbol{x} \in \overline{\Omega}. \end{aligned} \tag{8.31}$$

In order to apply a convolution operator to (8.30), one has to extend first all functions outside the domain. These functions will satisfy the Navier–Stokes equations in a distributional sense. Then, the convolution operator can be applied, filtering and differentiation commute, and the space-averaged Navier–Stokes equations are obtained. □

Lemma 8.36 (Regularity of the Normal Stress and Estimate of Its $L^p(\partial\Omega)$ Norm) *If* (8.31) *holds, then* $\mathbb{S}\boldsymbol{n} \in H^{1/2}(\partial\Omega)$. *In particular, for almost every* $t \in [0,T]$, $\mathbb{S}\boldsymbol{n} \in L^q(\partial\Omega)$ *with*

$$q \in [1,\infty) \text{ if } d = 2, \quad q \in [1,4] \text{ if } d = 3.$$

Proof By the trace theorem. Theorem A.34 and (A.11) for $s = 1$, it follows that $\nabla\boldsymbol{u} \in H^{1/2}(\partial\Omega)$ and $p \in H^{1/2}(\partial\Omega)$, such that $\mathbb{S}\boldsymbol{n}\ H^{1/2}(\partial\Omega)$. Theorem A.34 for $q = 2$ provides also the second statement of the lemma. ■

Remark 8.37 (Extensions of the Functions to $\mathbb{R}^d$) For deriving an equation to which the convolution with a filter function can be applied, $\boldsymbol{f}$ has to be extended off Ω and then $(\boldsymbol{u},p)$ must be extended compatibly with the extension of $\boldsymbol{f}$ such that the extended functions satisfy a kind of Navier–Stokes equations. Since the right-hand side is data, the extension of $\boldsymbol{f}$ has to be known. For $\overline{\boldsymbol{f}}$ to be easily to compute, $\boldsymbol{f}$ is extended by $\boldsymbol{0}$ off Ω. Then, $(\boldsymbol{u},p)$ can be extended by zero off Ω, too. This extension is reasonable since $\boldsymbol{u} = \boldsymbol{0}$ on $\partial\Omega$. Thus, one defines

$$\boldsymbol{u} = \boldsymbol{0}, \quad \boldsymbol{u}_0 = \boldsymbol{0}, \quad p = 0, \quad \boldsymbol{f} = \boldsymbol{0} \quad \text{if } \boldsymbol{x} \notin \overline{\Omega}.$$

The extended functions possess the following regularities

$$\begin{aligned} \boldsymbol{u} \in H_0^1\left(\mathbb{R}^d\right), \ p \in L_0^2\left(\mathbb{R}^d\right) \text{ for a.e. } t \in [0,T], \\ \boldsymbol{u} \in H^1((0,T)) \text{ for a.e. } \boldsymbol{x} \in \mathbb{R}^d. \end{aligned} \tag{8.32}$$

From (8.31) and (8.32), it follows that the first order weak derivatives of the extended velocity $\partial_t\boldsymbol{u}$, $\nabla\boldsymbol{u}$,$\nabla\cdot\boldsymbol{u}$, and $\nabla\cdot\left(\boldsymbol{u}\boldsymbol{u}^T\right)$ are well defined on $\mathbb{R}^d$, taking their indicated values in Ω and being identically zero off Ω.

An extension of $\boldsymbol{u}$ off Ω as a function in $H^2\left(\mathbb{R}^d\right)$ exists but it is unknown, in particular since $\boldsymbol{u}$ is not known away from the boundary. Using this extension, instead of $\boldsymbol{u} \equiv \boldsymbol{0}$ on $\mathbb{R}^d \setminus \Omega$, would lead to an unknown extension of $\boldsymbol{f}$ and hence $\overline{\boldsymbol{f}}$ would not be known in the space-averaged momentum equation. □

Remark 8.38 (Extension of the Pressure Term and the Viscous Term) Since $\boldsymbol{u} \notin H^2\left(\mathbb{R}^d\right)$, $p \notin H^1\left(\mathbb{R}^d\right)$, the terms $\nabla\cdot\mathbb{D}(\boldsymbol{u})$ and ∇p must be defined in the sense of distributions. To this end, let $\varphi \in C_0^\infty\left(\mathbb{R}^d\right)$. Since $p \equiv 0$ on $\mathbb{R}^d \setminus \Omega$ for all times, one obtains with the definition of the distributional derivative and with integration by parts

$$\begin{aligned} (\partial_i p)(\varphi)(t) &= -\int_{\mathbb{R}^d} p(t,\boldsymbol{x})\,\partial_i\varphi(\boldsymbol{x})\ d\boldsymbol{x} \\ &= \int_\Omega \varphi(\boldsymbol{x})\,\partial_i p(t,\boldsymbol{x})\ d\boldsymbol{x} - \int_{\partial\Omega} p(t,\boldsymbol{s})\,\varphi(\boldsymbol{s})\,\boldsymbol{e}_i\cdot\boldsymbol{n}(\boldsymbol{s})\ ds, \end{aligned}$$

$i = 1, \ldots, d$, or in compact form

$$(\nabla p)(\varphi)(t) = -\int_{\mathbb{R}^d} p(t, \boldsymbol{x}) \nabla\varphi(\boldsymbol{x}) \, d\boldsymbol{x} \quad (8.33)$$
$$= \int_{\Omega} \varphi(\boldsymbol{x}) \nabla p(t, \boldsymbol{x}) \, d\boldsymbol{x} - \int_{\partial\Omega} p(t, \boldsymbol{s}) \varphi(\boldsymbol{s}) \boldsymbol{n}(\boldsymbol{s}) \, d\boldsymbol{s}.$$

In the same way, one derives

$$\nabla \cdot \mathbb{D}(\boldsymbol{u})(\varphi)(t)$$
$$= -\int_{\mathbb{R}^d} \mathbb{D}(\boldsymbol{u})(t, \boldsymbol{x}) \nabla\varphi(\boldsymbol{x}) \, d\boldsymbol{x} \quad (8.34)$$
$$= \int_{\Omega} \varphi(\boldsymbol{x}) \nabla \cdot \mathbb{D}(\boldsymbol{u})(t, \boldsymbol{x}) \, d\boldsymbol{x} - \int_{\partial\Omega} \varphi(\boldsymbol{s}) \mathbb{D}(\boldsymbol{u})(t, \boldsymbol{s}) \boldsymbol{n}(\boldsymbol{s}) \, d\boldsymbol{s}.$$

Both distributions define continuous linear functionals on $C_0^\infty(\mathbb{R}^d)$ such that they have compact support. Note that from the regularity assumption (8.31), it follows that the traces of p and $\mathbb{D}(\boldsymbol{u})$ on $\partial\Omega$ are well defined. □

Remark 8.39 (The Distributional Form of the Momentum Equation) From (8.33) and (8.34), it follows that the extended functions $(\boldsymbol{u}, p)$ satisfy the following distributional form of the momentum equation

$$\left(\partial_t \boldsymbol{u} - 2\nu\nabla \cdot \mathbb{D}(\boldsymbol{u}) + \nabla \cdot \left(\boldsymbol{u}\boldsymbol{u}^T\right) + \nabla p\right)(\varphi)(t) \quad (8.35)$$
$$= \int_{\Omega} \boldsymbol{f}(t, \boldsymbol{x})\varphi(\boldsymbol{x}) \, d\boldsymbol{x} + \int_{\partial\Omega} \mathbb{S}(t, \boldsymbol{s}) \boldsymbol{n}(\boldsymbol{s}) \varphi(\boldsymbol{s}) \, d\boldsymbol{s}, \quad \forall \varphi \in C_0^\infty\left(\mathbb{R}^d\right).$$

□

Remark 8.40 (The Space-Averaged Navier–Stokes Equations in a Bounded Domain) The space-averaged Navier–Stokes equations are now derived by convolving (8.35) with a filter function $g_{\text{fil}}(\boldsymbol{x}) \in C^\infty\left(\mathbb{R}^d\right)$. Applying the convolution with g_{fil} to (8.35) gives a function in $C^\infty\left(\mathbb{R}^d\right)$ and moreover convolution and differentiation commute on $\mathbb{R}^d$, see Hörmander (1990, Theorem 4.1.1). One obtains for the left-hand side of (8.35)

$$g_{\text{fil}} * \left[\left(\partial_t \boldsymbol{u} - 2\nu\nabla \cdot \mathbb{D}(\boldsymbol{u}) + \nabla \cdot \left(\boldsymbol{u}\boldsymbol{u}^T\right) + \nabla p\right)(\varphi)\right]$$
$$= \partial_t \overline{\boldsymbol{u}} - 2\nu\nabla \cdot \mathbb{D}(\overline{\boldsymbol{u}}) + \nabla \cdot \left(\overline{\boldsymbol{u}\boldsymbol{u}^T}\right) + \nabla \overline{p} \quad (8.36)$$

in $(0, T] \times \mathbb{R}^d$.

The filter of the viscous term and the pressure term will be studied in more detail. Let $H(\varphi)$ be a distribution with compact support that has the form

$$H(\varphi) = -\int_{\mathbb{R}^d} f(\boldsymbol{x})\, D^{\boldsymbol{\alpha}}\varphi(\boldsymbol{x})\ d\boldsymbol{x},$$

where $D^{\boldsymbol{\alpha}}$ is the derivative of φ with the multi-index $\boldsymbol{\alpha}$. Then, $H * g_{\text{fil}} \in C^\infty(\mathbb{R}^d)$, see Rudin (1991, Theorem 6.35), where

$$\overline{H}(\boldsymbol{x}) = (H * g_{\text{fil}})(\boldsymbol{x}) = H(g_{\text{fil}}(\boldsymbol{x}-\cdot)) = -\int_{\mathbb{R}^d} f(\boldsymbol{y})\, D^{\boldsymbol{\alpha}} g_{\text{fil}}(\boldsymbol{x}-\boldsymbol{y})\ d\boldsymbol{y}. \tag{8.37}$$

Applying (8.37) to (8.33) yields

$$\begin{aligned} \nabla\overline{p}\,(t,\boldsymbol{x}) &= g_{\text{fil}} * ((\nabla p)(\varphi))(t,\boldsymbol{x}) \\ &= -\int_{\mathbb{R}^d} p(t,\boldsymbol{y})\,\nabla g_{\text{fil}}(\boldsymbol{x}-\boldsymbol{y})\ d\boldsymbol{y} \qquad (8.38) \\ &= \int_{\Omega} \nabla p(t,\boldsymbol{y})\, g_{\text{fil}}(\boldsymbol{x}-\boldsymbol{y})\ d\boldsymbol{y} - \int_{\partial\Omega} g_{\text{fil}}(\boldsymbol{x}-\boldsymbol{s})\, p(t,\boldsymbol{s})\,\boldsymbol{n}(\boldsymbol{s})\ d\boldsymbol{s}. \end{aligned}$$

Convolving (8.34) in the same way gives

$$\begin{aligned} &\nabla\cdot\mathbb{D}(\overline{\boldsymbol{u}})(t,\boldsymbol{x}) \qquad (8.39) \\ &= \int_{\Omega} \nabla\cdot\mathbb{D}(\boldsymbol{u})(t,\boldsymbol{y})\, g_{\text{fil}}(\boldsymbol{x}-\boldsymbol{y})\ d\boldsymbol{y} - \int_{\partial\Omega} g_{\text{fil}}(\boldsymbol{x}-\boldsymbol{s})\,\mathbb{D}(\boldsymbol{u})(t,\boldsymbol{s})\,\boldsymbol{n}(\boldsymbol{s})\ d\boldsymbol{s}. \end{aligned}$$

Combining (8.36), (8.38), and (8.39) leads to the space-averaged momentum equation

$$\begin{aligned} &\partial_t\overline{\boldsymbol{u}} - 2\nu\nabla\cdot\mathbb{D}(\overline{\boldsymbol{u}}) + \nabla\cdot\left(\overline{\boldsymbol{u}\boldsymbol{u}^T}\right) + \nabla\overline{p} \\ &= \overline{\boldsymbol{f}} + \int_{\partial\Omega} g_{\text{fil}}(\boldsymbol{x}-\boldsymbol{s})\,\mathbb{S}(t,\boldsymbol{s})\,\boldsymbol{n}(\boldsymbol{s})\ d\boldsymbol{s} \quad \text{in } (0,T]\times\mathbb{R}^d \qquad (8.40) \end{aligned}$$

with

$$\overline{\boldsymbol{f}}(t,\boldsymbol{x}) = \int_{\Omega} \left(\partial_t\boldsymbol{u} - 2\nu\nabla\cdot\mathbb{D}(\boldsymbol{u}) + \nabla\cdot\left(\boldsymbol{u}\boldsymbol{u}^T\right) + \nabla p\right)(t,\boldsymbol{y}) g_{\text{fil}}(\boldsymbol{x}-\boldsymbol{y})\ d\boldsymbol{y}. \tag{8.41}$$

Since $\boldsymbol{f}$ vanishes outside Ω for $t \in (0,T]$, (8.41) has the same form which is obtained if $\boldsymbol{f}$ is filtered directly. This correspondence has to be expected from a consistent extension of the functions. □

Definition 8.41 (Commutation Error) Let $g_{\mathrm{fil}} \in C^{\infty}\left(\mathbb{R}^d\right)$ be a filter function with filter width δ. The commutation error $A_\delta\left(\mathbb{S}\right)$ in the space-averaged Navier–Stokes equations is defined to be

$$A_\delta\left(\mathbb{S}\right)(t,\boldsymbol{x}) = \int_{\partial\Omega} g_{\mathrm{fil}}\left(\boldsymbol{x}-\boldsymbol{s}\right)\mathbb{S}\left(t,\boldsymbol{s}\right)\boldsymbol{n}\left(\boldsymbol{s}\right)\, d\boldsymbol{s}.$$

□

Remark 8.42 (On the Commutation Error)

- The analysis of the commutation error will be performed for the Gaussian filter g_{Gauss} defined in Example 8.25.
- In the analysis of the commutation error, an arbitrary but fixed time $t \in (0,T]$ will be considered such that the dependency of $A_\delta\left(\mathbb{S}\right)$ on the time can be neglected.
- If the viscous term in the Navier–Stokes equations is written as $\nu\Delta\boldsymbol{u}$ instead of $2\nu\nabla\cdot\mathbb{D}\left(\boldsymbol{u}\right)$, the resulting space-averaged momentum equation is given by replacing $2\nu\mathbb{D}\left(\overline{\boldsymbol{u}}\right)$ in (8.40) by $\nu\nabla\overline{\boldsymbol{u}}$ and $2\nu\mathbb{D}\left(\boldsymbol{u}\right)(t,\boldsymbol{s})$ by $\nu\nabla\boldsymbol{u}\left(t,\boldsymbol{s}\right)$ in the stress tensor.
- The space-averaged Navier–Stokes equations arising from the Navier–Stokes equations in a bounded domain thus possess the extra boundary integral, $A_\delta\left(\mathbb{S}\right)$. Omitting this integral results in committing a commutation error. Including this integral in the space-averaged momentum equation introduces a new modeling question since it depends on the unknown normal stress on $\partial\Omega$ of $(\boldsymbol{u},p)$ but not of $(\overline{\boldsymbol{u}},\overline{p})$.

□

8.2.4 Analysis of the Commutation Error for the Gaussian Filter

Remark 8.43 (The Term to be Studied) In view of Definition 8.41, Lemma 8.36, and Remark 8.42, one has to study terms of the form

$$\int_{\partial\Omega} g_{\mathrm{Gauss}}\left(\boldsymbol{x}-\boldsymbol{s}\right)\psi\left(\boldsymbol{s}\right) d\boldsymbol{s} \tag{8.42}$$

with $\psi \in L^q\left(\partial\Omega\right)$, $1 \le q \le \infty$. Recall that $\partial\Omega$ is assumed to possess a finite $(d-1)$-dimensional measure. □

Lemma 8.44 (Expression (8.42) Defines a Function in $L^p\left(\mathbb{R}^d\right)$, $p \in [1,\infty]$) *Let $\psi \in L^q\left(\partial\Omega\right)$, $1 \le q \le \infty$, then (8.42) belongs to $L^p\left(\mathbb{R}^d\right)$, $1 \le p \le \infty$.*

Proof

i) $p = \infty, q > 1$. By the Hölder inequality (A.9), one obtains with $r^{-1} + q^{-1} = 1$, $q > 1$,

$$\begin{aligned}
&\left| \int_{\partial\Omega} g_{\mathrm{Gauss}}\left(\boldsymbol{x}-\boldsymbol{s}\right) \psi\left(\boldsymbol{s}\right) \, d\boldsymbol{s} \right| \\
&\leq \left(\int_{\partial\Omega} g_{\mathrm{Gauss}}^{r}\left(\boldsymbol{x}-\boldsymbol{s}\right) \, d\boldsymbol{s} \right)^{1/r} \|\psi\|_{L^{q}(\partial\Omega)} \\
&= \left(\int_{\partial\Omega} \left(\frac{6}{\delta^2\pi}\right)^{rd/2} \exp\left(-\frac{6r}{\delta^2}\|\boldsymbol{x}-\boldsymbol{s}\|_2^2\right) \, d\boldsymbol{s} \right)^{1/r} \|\psi\|_{L^{q}(\partial\Omega)}. \qquad (8.43)
\end{aligned}$$

By the triangle inequality and Young's inequality (A.5), it follows that

$$\|\boldsymbol{x}\|_2^2 \leq 2\|\boldsymbol{x}-\boldsymbol{s}\|_2^2 + 2\|\boldsymbol{s}\|_2^2 \quad \Longleftrightarrow \quad 2\|\boldsymbol{x}-\boldsymbol{s}\|_2^2 \geq \|\boldsymbol{x}\|_2^2 - 2\|\boldsymbol{s}\|_2^2,$$

which implies, together with the monotonicity of the exponential, that

$$\exp\left(-\frac{6r\|\boldsymbol{x}-\boldsymbol{s}\|_2^2}{\delta^2}\right) \leq \exp\left(3r\frac{-\|\boldsymbol{x}\|_2^2 + 2\|\boldsymbol{s}\|_2^2}{\delta^2}\right).$$

Inserting this estimate in (8.43) yields

$$\begin{aligned}
&\left| \int_{\partial\Omega} g_{\mathrm{Gauss}}\left(\boldsymbol{x}-\boldsymbol{s}\right) \psi\left(\boldsymbol{s}\right) \, d\boldsymbol{s} \right| \\
&\leq \left(\frac{6}{\delta^2\pi}\right)^{d/2} \|\psi\|_{L^{q}(\partial\Omega)} \left(\int_{\partial\Omega} \exp\left(\frac{6r\|\boldsymbol{s}\|_2^2}{\delta^2}\right) \, d\boldsymbol{s} \right)^{1/r} \exp\left(-\frac{3\|\boldsymbol{x}\|_2^2}{\delta^2}\right) \\
&< \infty, \qquad (8.44)
\end{aligned}$$

since $\partial\Omega$ has finite measure in $\mathbb{R}^{d-1}$ and the exponential is a bounded function on $\partial\Omega$. This estimate proves the statement for $L^{\infty}\left(\mathbb{R}^d\right)$ and $q > 1$.

ii) $p \in [1, \infty), q > 1$. The proof for $p \in [1, \infty)$ and $q > 1$ is obtained by raising both sides of (8.44) to the power p, integrating on $\mathbb{R}^d$, and using (8.16), from what follows that

$$\int_{\mathbb{R}^d} \exp\left(-\frac{3p\|\boldsymbol{x}\|_2^2}{\delta^2}\right) \, d\boldsymbol{x} < \infty.$$

iii) $p \in [1, \infty), q = 1$. If $q = 1$, one has for $1 \leq p < \infty$ with Hölder's inequality (A.9) and the monotonicity of the Gaussian filter (8.23), such that the largest values are taken if the absolute value of the argument is as small is possible,

$$\int_{\mathbb{R}^d} \left| \int_{\partial\Omega} g_{\text{Gauss}}(\boldsymbol{x}-\boldsymbol{s})\, \psi(\boldsymbol{s})\, d\boldsymbol{s} \right|^p d\boldsymbol{x} \leq \int_{\mathbb{R}^d} \sup_{\boldsymbol{s}\in\partial\Omega} g^p_{\text{Gauss}}(\boldsymbol{x}-\boldsymbol{s})\, d\boldsymbol{x} \left(\int_{\partial\Omega} \psi(\boldsymbol{s})\, d\boldsymbol{s} \right)^p$$
$$= \int_{\mathbb{R}^d} g^p_{\text{Gauss}}(d(\boldsymbol{x}, \partial\Omega))\, d\boldsymbol{x}\, \|\psi\|^p_{L^1(\partial\Omega)}.$$

Choose a ball $B(\mathbf{0}, R)$ with sufficiently large radius R such that $d(\boldsymbol{x}, \partial\Omega) > \|\boldsymbol{x}\|_2/2$ for all $\boldsymbol{x} \notin B(\mathbf{0}, R)$. Then, the integral on $\mathbb{R}^d$ is split into a sum of two integrals. The first integral is computed on $B(\mathbf{0}, R)$. It is finite since the term in the integral is a continuous function on $\overline{B}(\mathbf{0}, R)$. The second integral on $\mathbb{R}^d \setminus B(\mathbf{0}, R)$ is also finite because

$$\int_{\mathbb{R}^d\setminus B(\mathbf{0},R)} g^p_{\text{Gauss}}(d(\boldsymbol{x}, \partial\Omega))\, d\boldsymbol{x} \leq \int_{\mathbb{R}^d} g^p_{\text{Gauss}}\left(\frac{\|\boldsymbol{x}\|_2}{2} \right) d\boldsymbol{x}$$

and the integrability of the Gaussian filter (8.21). Hence, the statement is proved for $p < \infty$.

iv) $p = \infty, q = 1$. For $p = \infty$ and $q = 1$, one finds

$$\operatorname*{ess\,sup}_{\boldsymbol{x}\in\mathbb{R}^d} \left| \int_{\partial\Omega} g_{\text{Gauss}}(\boldsymbol{x}-\boldsymbol{s})\, \psi(\boldsymbol{s})\, d\boldsymbol{s} \right| \leq \operatorname*{ess\,sup}_{\boldsymbol{x}\in\mathbb{R}^d} \operatorname*{ess\,sup}_{\boldsymbol{s}\in\partial\Omega} g_{\text{Gauss}}(\boldsymbol{x}-\boldsymbol{s})\, \|\psi\|_{L^1(\partial\Omega)}$$
$$\leq g_{\text{Gauss}}(\mathbf{0})\, \|\psi\|_{L^1(\partial\Omega)} < \infty,$$

where in the second estimate, it was used that the Gaussian filter takes its largest value at the origin, which follows from its monotonicity (8.23). This estimate finishes the proof of the lemma. ■

Theorem 8.45 (Behavior of the $L^p(\mathbb{R}^d)$ Norm of the Commutation Error for $\delta \to 0$) *Let $\psi \in L^p(\partial\Omega)$, $1 \leq p \leq \infty$. A necessary and sufficient condition for*

$$\lim_{\delta\to 0} \left\| \int_{\partial\Omega} g_{\text{Gauss}}(\boldsymbol{x}-\boldsymbol{s})\, \psi(\boldsymbol{s})\, d\boldsymbol{s} \right\|_{L^p(\mathbb{R}^d)} = 0, \tag{8.45}$$

$1 \leq p \leq \infty$, is that ψ vanishes almost everywhere on $\partial\Omega$.

Proof It is obvious that the condition is sufficient.

Let (8.45) hold. From Hölder's inequality (A.9) and (8.24), one obtains for an arbitrary function $\varphi \in C_0^\infty(\mathbb{R}^d)$

$$\lim_{\delta\to 0} \left| \int_{\mathbb{R}^d} \varphi(\boldsymbol{x}) \left(\int_{\partial\Omega} g_{\text{Gauss}}(\boldsymbol{x}-\boldsymbol{s})\, \psi(\boldsymbol{s})\, d\boldsymbol{s} \right) d\boldsymbol{x} \right|$$
$$\leq \lim_{\delta\to 0} \|\varphi\|_{L^q(\mathbb{R}^d)} \left\| \int_{\partial\Omega} g_{\text{Gauss}}(\boldsymbol{x}-\boldsymbol{s})\, \psi(\boldsymbol{s})\, d\boldsymbol{s} \right\|_{L^p(\mathbb{R}^d)} = 0, \tag{8.46}$$

where $p^{-1} + q^{-1} = 1$. By Fubini's theorem and the symmetry of the Gaussian filter (8.22), one gets

$$\begin{aligned}
&\lim_{\delta\to 0}\int_{\mathbb{R}^d}\varphi(\boldsymbol{x})\left(\int_{\partial\Omega} g_{\text{Gauss}}(\boldsymbol{x}-\boldsymbol{s})\,\psi(\boldsymbol{s})\,\,d\boldsymbol{s}\right)\,d\boldsymbol{x}\\
&=\lim_{\delta\to 0}\int_{\partial\Omega}\psi(\boldsymbol{s})\left(\int_{\mathbb{R}^d} g_{\text{Gauss}}(\boldsymbol{x}-\boldsymbol{s})\,\varphi(\boldsymbol{x})\,\,d\boldsymbol{x}\right)\,d\boldsymbol{s}\\
&=\int_{\partial\Omega}\psi(\boldsymbol{s})\lim_{\delta\to 0}\left(\int_{\mathbb{R}^d} g_{\text{Gauss}}(\boldsymbol{s}-\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\,d\boldsymbol{x}\right)\,d\boldsymbol{s}=\int_{\partial\Omega}\psi(\boldsymbol{s})\,\varphi(\boldsymbol{s})\,\,d\boldsymbol{s},
\end{aligned}$$

where the last step is a consequence of Lemma 8.26 ii) since $\varphi \in L^\infty\left(\mathbb{R}^d\right)$ and φ is uniformly continuous on the compact set $\partial\Omega$. Thus, with (8.46), it follows that

$$0=\int_{\mathbb{R}^d}\varphi(\boldsymbol{x})\left(\int_{\partial\Omega} g_{\text{Gauss}}(\boldsymbol{x}-\boldsymbol{s})\,\psi(\boldsymbol{s})\,\,d\boldsymbol{s}\right)\,d\boldsymbol{x}=\int_{\partial\Omega}\psi(\boldsymbol{s})\,\varphi(\boldsymbol{s})\,\,d\boldsymbol{s}\quad\forall\,\varphi\in C_0^\infty\left(\mathbb{R}^d\right).$$

This statement is true if and only if ψ vanishes almost everywhere on $\partial\Omega$. ■

Remark 8.46 (Interpretation of Lemma 8.44 and Theorem 8.45) Lemma 8.44 states that the commutation error $A_\delta(\mathbb{S})$ is a function in $L^p\left(\mathbb{R}^d\right)$, $1 \le p \le \infty$. Then, it is shown in Theorem 8.44 that $A_\delta(\mathbb{S})$ vanishes in $L^p\left(\mathbb{R}^d\right)$ as $\delta \to 0$ if and only if the normal stress is identical to zero almost everywhere on $\partial\Omega$. This condition means that the wall has zero influence on the wall-bounded turbulent flow. This situation is not expected to be satisfied in any interesting flow problem ! If the commutation error term is simply dropped and then the strong form of the space-averaged Navier–Stokes equations is discretized, as, e.g., by a finite difference method, this result shows that the committed commutation error is $\mathcal{O}(1)$! □

Remark 8.47 (Further Results on the $L^p\left(\mathbb{R}^d\right)$ Norm of the Commutation Error) With a quite technical proof, a bound for the $L^p\left(\mathbb{R}^d\right)$ norm of the commutation error in terms of δ can be derived, see Dunca et al. (2004) or John (2004, Theorem 3.13). An inspection of the proof shows that the commutation error is largest at the boundary and it decays away from the boundary. □

Remark 8.48 (Motivation for Considering the $H^{-1}(\Omega)$ Norm of the Commutation Error) Variational methods, such as finite element methods, discretize the weak form of the considered equations. For these methods, the $H^{-1}(\Omega)$ norm of any omitted term is of interest. □

Lemma 8.49 (Estimate of $\|\overline{v}-v\|_{H^{1/2}(\mathbb{R}^d)}$) *There exists a constant C which does not depend on v and δ such that*

$$\|\overline{v}-v\|_{H^{1/2}(\mathbb{R}^d)} \le C\delta^{1/2}\,\|v\|_{H^1(\mathbb{R}^d)} \tag{8.47}$$

for any $v \in H^1\left(\mathbb{R}^d\right)$ and any $\delta > 0$.

Proof The norm in $H^{1/2}\left(\mathbb{R}^d\right)$ can be expressed with the Fourier transform, e.g., see Evans (2010, Sect. 5.8.5, Theorem 8),

$$\|v\|^2_{H^{1/2}\left(\mathbb{R}^d\right)} = \int_{\mathbb{R}^d} \left(1 + \|\boldsymbol{x}\|_2^{1/2}\right)^2 |\mathcal{F}(v)|^2 \, d\boldsymbol{x}.$$

Applying Young's inequality, it follows that

$$\left(1 + \|\boldsymbol{x}\|_2^2\right)^{1/2} \le \left(1 + \|\boldsymbol{x}\|_2^{1/2}\right)^2 \le C\left(1 + \|\boldsymbol{x}\|_2^2\right)^{1/2},$$

such that

$$\|v\|^2_{H^{1/2}\left(\mathbb{R}^d\right)} = \int_{\mathbb{R}^d} \left(1 + \|\boldsymbol{x}\|_2^2\right)^{1/2} |\mathcal{F}(v)|^2 \, d\boldsymbol{x}$$

is an equivalent norm.

Using the latter definition, the linearity of the Fourier transform, and (8.13) yields

$$\begin{aligned}
\|\,\overline{v} - v\|^2_{H^{1/2}\left(\mathbb{R}^d\right)} &= \int_{\mathbb{R}^d} \left(1 + \|\boldsymbol{x}\|_2^2\right)^{1/2} |\mathcal{F}(g_{\text{Gauss}} * v - v)|^2 \, d\boldsymbol{x} \\
&= \int_{\mathbb{R}^d} \left(1 + \|\boldsymbol{x}\|_2^2\right)^{1/2} |\mathcal{F}(g_{\text{Gauss}} * v) - \mathcal{F}(v)|^2 \, d\boldsymbol{x} \\
&= \int_{\mathbb{R}^d} \left(1 + \|\boldsymbol{x}\|_2^2\right)^{1/2} |\mathcal{F}(g_{\text{Gauss}})\,\mathcal{F}(v) - \mathcal{F}(v)|^2 \, d\boldsymbol{x} \\
&= \int_{\mathbb{R}^d} \left(1 + \|\boldsymbol{x}\|_2^2\right)^{1/2} |1 - \mathcal{F}(g_{\text{Gauss}})|^2 \, |\mathcal{F}(v)|^2 \, d\boldsymbol{x} \\
&= \int_{\{\|\boldsymbol{x}\|_2 > \pi/\delta\}} \left(1 + \|\boldsymbol{x}\|_2^2\right)^{1/2} |1 - \mathcal{F}(g_{\text{Gauss}})|^2 \, |\mathcal{F}(v)|^2 \, d\boldsymbol{x} \\
&\quad + \int_{\{\|\boldsymbol{x}\|_2 \le \pi/\delta\}} \left(1 + \|\boldsymbol{x}\|_2^2\right)^{1/2} |1 - \mathcal{F}(g_{\text{Gauss}})|^2 \, |\mathcal{F}(v)|^2 \, d\boldsymbol{x}.
\end{aligned}$$

The integrals on the right-hand side will be bounded separately.

The bound of the first integral relies on the fact that $\|\boldsymbol{x}\|_2$ is sufficiently large. There exists a constant $C > 0$, which does not depend on δ and v, such that for $\|\boldsymbol{x}\|_2 > \pi/\delta$

$$\frac{1}{\left(1 + \|\boldsymbol{x}\|_2^2\right)^{1/2}} < \frac{1}{\left(1 + \pi^2/\delta^2\right)^{1/2}} = \delta \frac{1}{\left(\delta^2 + \pi^2\right)^{1/2}} \le \frac{\delta}{\pi^{1/2}} = C\delta.$$

From (8.17), it follows the pointwise estimate

$$|1 - \mathcal{F}(g_{\text{Gauss}})(\boldsymbol{x})| \le 1 \quad \forall\, \boldsymbol{x} \in \mathbb{R}^d.$$

Thus, the first integral can be bounded by

$$
\begin{aligned}
&\left| \int_{\{\|\boldsymbol{x}\|_2 > \pi/\delta\}} \left(1 + \|\boldsymbol{x}\|_2^2\right)^{1/2} |1 - \mathcal{F}(g_{\mathrm{Gauss}})|^2 \, |\mathcal{F}(v)|^2 \, d\boldsymbol{x} \right| \\
&\le \int_{\{\|\boldsymbol{x}\|_2 > \pi/\delta\}} \left(1 + \|\boldsymbol{x}\|_2^2\right) \left(1 + \|\boldsymbol{x}\|_2^2\right)^{-1/2} |\mathcal{F}(v)|^2 \, d\boldsymbol{x} \\
&\le C\delta \int_{\{\|\boldsymbol{x}\|_2 > \pi/\delta\}} \left(1 + \|\boldsymbol{x}\|_2^2\right) |\mathcal{F}(v)|^2 \, d\boldsymbol{x}.
\end{aligned} \tag{8.48}
$$

To bound the second integral, it is used that the Fourier transform of the Gaussian filter is close to one at the origin. Applying a Taylor series expansion of (8.17) at $\boldsymbol{x} = \mathbf{0}$ for fixed δ yields

$$
\mathcal{F}(g_{\mathrm{Gauss}})(\boldsymbol{x}) = 1 - \frac{\delta^2 \|\boldsymbol{x}\|_2^2}{24} + \mathcal{O}\left(\delta^4 \|\boldsymbol{x}\|_2^4\right).
$$

One obtains for all $\boldsymbol{x}$ with $\|\boldsymbol{x}\|_2 \le \pi/\delta$ the pointwise bound

$$
|1 - \mathcal{F}(g_{\mathrm{Gauss}})(\boldsymbol{x})|^2 \le C\delta^4 \|\boldsymbol{x}\|_2^4 \le C\delta^4 \frac{\pi^3}{\delta^3} \|\boldsymbol{x}\|_2 = C\delta \|\boldsymbol{x}\|_2 ,
$$

where C does not depend on δ or $\boldsymbol{x}$. Continuing this estimate with $\|\boldsymbol{x}\|_2 \le \left(1 + \|\boldsymbol{x}\|_2^2\right)^{1/2}$ shows that the second integral can be bounded as follows

$$
\begin{aligned}
&\left| \int_{\{\|\boldsymbol{x}\|_2 \le \pi/\delta\}} \left(1 + \|\boldsymbol{x}\|_2^2\right)^{1/2} |1 - \mathcal{F}(g_{\mathrm{Gauss}})|^2 \, |\mathcal{F}(v)|^2 \, d\boldsymbol{x} \right| \\
&\le C\delta \int_{\{\|\boldsymbol{x}\|_2 \le \pi/\delta\}} \left(1 + \|\boldsymbol{x}\|_2^2\right) |\mathcal{F}(v)|^2 \, d\boldsymbol{x}.
\end{aligned} \tag{8.49}
$$

Combining (8.48) and (8.49) gives

$$
\| \overline{v} - v\|_{H^{1/2}(\mathbb{R}^d)}^2 \le C\delta \int_{\mathbb{R}^d} \left(1 + \|\boldsymbol{x}\|_2^2\right) |\mathcal{F}(v)|^2 \, d\boldsymbol{x} = C\delta \|v\|_{H^1(\mathbb{R}^d)}^2 ,
$$

which is the statement of the lemma. ■

Theorem 8.50 (Convergence of the Commutation Error in $H^{-1}(\Omega)$) *Let $\psi \in L^2(\partial\Omega)$, then there exists a constant $C > 0$, which depends only on Ω, such that*

$$
\left\| \int_{\partial\Omega} g_{\mathrm{Gauss}}(\boldsymbol{x} - \boldsymbol{s}) \, \psi(\boldsymbol{s}) \, d\boldsymbol{s} \right\|_{H^{-1}(\Omega)} \le C\delta^{1/2} \|\psi\|_{L^2(\partial\Omega)} \tag{8.50}
$$

for every $\delta > 0$.

Proof It is

$$\|\varphi\|_{H^{-1}(\Omega)} = \sup_{v \in H_0^1(\Omega)} \frac{\int_\Omega (v\varphi)(\boldsymbol{x})\, d\boldsymbol{x}}{\|\nabla v\|_{L^2(\Omega)}}. \tag{8.51}$$

The numerator of the right-hand side will be estimated for the commutation error.

Let $v \in H_0^1(\Omega)$. Extending v by zero outside Ω, applying Fubini's theorem, utilizing the symmetry of the Gaussian filter (8.22), using that v vanishes on $\partial\Omega$, applying the Cauchy–Schwarz inequality (A.10), the trace theorem (A.11) for $s = 1/2$, Lemma 8.49, that v vanishes off Ω, and the Poincaré inequality (A.12) yields

$$\begin{aligned}
&\int_\Omega \left(\int_{\partial\Omega} g_{\text{Gauss}}(\boldsymbol{x}-\boldsymbol{s})\,\psi(\boldsymbol{s})\, d\boldsymbol{s}\right) v(\boldsymbol{x})\, d\boldsymbol{x} \\
&= \int_{\partial\Omega} \psi(\boldsymbol{s}) \left(\int_\Omega g_{\text{Gauss}}(\boldsymbol{x}-\boldsymbol{s})\, v(\boldsymbol{x})\, d\boldsymbol{x}\right) d\boldsymbol{s} \\
&= \int_{\partial\Omega} \psi(\boldsymbol{s}) \left(\int_\Omega g_{\text{Gauss}}(\boldsymbol{s}-\boldsymbol{x})\, v(\boldsymbol{x})\, d\boldsymbol{x}\right) d\boldsymbol{s} = \int_{\partial\Omega} \psi(\boldsymbol{s})\, \overline{v}(\boldsymbol{s})\, d\boldsymbol{s} \\
&= \int_{\partial\Omega} \psi(\boldsymbol{s}) \left(\overline{v}(\boldsymbol{s}) - v(\boldsymbol{s})\right) d\boldsymbol{s} \le \|\overline{v} - v\|_{L^2(\partial\Omega)} \|\psi\|_{L^2(\partial\Omega)} \\
&\le C \|\overline{v} - v\|_{H^{1/2}(\Omega)} \|\psi\|_{L^2(\partial\Omega)} \le C \|\overline{v} - v\|_{H^{1/2}(\mathbb{R}^d)} \|\psi\|_{L^2(\partial\Omega)} \\
&\le C\delta^{1/2} \|v\|_{H^1(\mathbb{R}^d)} \|\psi\|_{L^2(\partial\Omega)} = C\delta^{1/2} \|v\|_{H^1(\Omega)} \|\psi\|_{L^2(\partial\Omega)} \\
&\le C\delta^{1/2} \|\nabla v\|_{L^2(\Omega)} \|\psi\|_{L^2(\partial\Omega)}.
\end{aligned}$$

Inserting this estimate in (8.51) gives the result of the theorem. ■

Remark 8.51 (Interpretation of Theorem 8.50) Theorem 8.50 shows that the commutation error tends to zero in $H^{-1}(\Omega)$ as $\delta \to 0$. The order of convergence is at least $\mathcal{O}\left(\delta^{1/2}\right)$. Thus, using variational methods, like the finite element method, leads to the expected asymptotic vanishing of the commutation error. It is not known whether the estimate (8.50) is optimal. □

Remark 8.52 (On a Weak Form of the Commutation Error) One can derive an estimate for a weak form of the commutation error, i.e., the commutation error is multiplied with a test function and integrated. Considering any $v \in H^1\left(\mathbb{R}^d\right)$ with $v|_\Omega \in H^2(\Omega) \cap V$ and $v(\boldsymbol{x}) = 0$ for $\boldsymbol{x} \notin \overline{\Omega}$, then the estimate is of the form

$$\begin{aligned}
&\int_{\mathbb{R}^d} \left| \overline{v}(\boldsymbol{x}) \int_{\partial\Omega} g_{\text{Gauss}}(\boldsymbol{x}-\boldsymbol{s})\,\psi(\boldsymbol{s})\, d\boldsymbol{s} \right|^k d\boldsymbol{x} \\
&\le C\delta^{1+\left(-d+\frac{(d-1)\theta}{q}+\beta\theta\right)k} \|\psi\|_{L^p(\partial\Omega)}^k \|v\|_{H^2(\Omega)}^k,
\end{aligned} \tag{8.52}$$

where $\delta \in (0, \epsilon)$, $\epsilon > 0$, $\theta \in (0, 1)$, $k \in [1, \infty)$, $\beta \in (0, 1)$ if $d = 2$ and $\beta = 1/2$ if $d = 3$, $p^{-1} + q^{-1} = 1$, $p > 1$, and C and ϵ depend on θ, k, and $|\partial\Omega|$. The proof can be found in Dunca et al. (2004) and John (2004, Sect. 3.7).

For $d = 2$, $k = 1$, and $p < \infty$ arbitrary large, i.e., ψ is sufficiently smooth, one finds that q is arbitrary close to 1. Choosing θ and β also arbitrary close to 1 leads to the following power of δ on the right-hand side of (8.52)

$$1 + (-2 + (1 - \epsilon_1) + (1 - \epsilon_2)) = 1 - (\epsilon_1 + \epsilon_2) = 1 - \epsilon_3$$

for arbitrary small $\epsilon_1, \epsilon_2, \epsilon_3 > 0$. In this case, the convergence is almost of first order.

Estimate (8.52) does not provide convergence for $d = 3$. Lemma 8.36 suggests choosing $p = 4$, i.e., $q = 4/3$. Then, for $k = 1$, the power of δ on the right-hand side of (8.52) becomes $2\,(\theta - 1)$, which is negative for $\theta < 1$. □

Remark 8.53 (Non-constant Filter Width and Skewed Filtering) The application of a filter is called skewed if the point in which a function is filtered and the center of the filter kernel do not coincide. Possible advantages of studying skewed filters are discussed in van der Bos and Geurts (2005). The skewed version of the Gaussian filter with skewness $\tilde{z}(x)$ and with variable filter width $\delta(x)$, in one dimension and neglecting the dependency on time, reads as follows

$$\overline{v}\ (x) = \frac{\sqrt{6}}{\delta(x)\pi} \int_{-\infty}^{\infty} \exp\left(-6\left(\frac{z + \tilde{z}(x)}{\delta(x)}\right)^2\right) v(x - z)\, dz.$$

Both, the variable filter width and the skewness introduce in general commutation errors. Estimates for the error $\left(\partial_i \overline{v} - \overline{\partial_i v}\right)(\boldsymbol{x})$ can be found in Berselli et al. (2007). □

8.2.5 Analysis of the Commutation Error for the Box Filter

Remark 8.54 (Motivation for Using the Box Filter) The box filter, see Example 8.27, is a filter whose kernel has a compact support. It will be required that the application of this filter leads to integrals whose domain of integration is a subset of $\overline{\Omega}$, i.e., the filter width at a point $\boldsymbol{x}$ in any direction is not allowed to be larger than the distance of $\boldsymbol{x}$ to the boundary in that direction. This situation has the appealing property that an extension of the function v to be filtered outside Ω is not necessary. Note that the non-smooth extension of the functions was the origin of the commutation error studied in Sect. 8.2.4. But the requirement that the domain of filtering is always in $\overline{\Omega}$ also implies that the filter width has to tend to zero (at least in one direction) as the point $\boldsymbol{x}$ in which v is filtered tends to the boundary of Ω. Thus, necessarily, the filter width is not constant but it is a function of $\boldsymbol{x}$. This property leads to a commutation error. □

Remark 8.55 (Normalized Box Filter) The filter kernel of the normalized box filter is given by

$$g_{\text{box}}(x) = \begin{cases} 1 & \text{for } x \in \left[-\frac{1}{2}, \frac{1}{2}\right], \\ 0 & \text{else.} \end{cases}$$

It follows for the first moments that

$$\int_{-1/2}^{1/2} g_{\text{box}}(x)\, dx = 1\,, \quad \int_{-1/2}^{1/2} g_{\text{box}}(x) x\, dx = 0\,, \quad \int_{-1/2}^{1/2} g_{\text{box}} x^2\, dx = \frac{1}{12}.$$

□

Remark 8.56 (The Box Filter with Variable Filter Width) Let $\Omega \subset \mathbb{R}^d$ be a bounded domain, $v \in C^1(\overline{\Omega})$, $\delta_l(\boldsymbol{x})$ be scalar filter width functions with $\delta_l(\boldsymbol{x}) \in C^1(\overline{\Omega})$, $\delta_l(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in \overline{\Omega}$, and $\delta_l(\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \in \Omega$, $l = 1, \ldots, d$. The support of the filter will be denoted by $B(\boldsymbol{x}) = [-\delta_1(\boldsymbol{x}), \delta_1(\boldsymbol{x})] \times \cdots \times [-\delta_d(\boldsymbol{x}), \delta_d(\boldsymbol{x})]$ and it is assumed that the filtering is applied for each point such that this support is in the closure of the domain, i.e., that

$$\boldsymbol{x} + B(\boldsymbol{x}) := [x_1 - \delta_1(\boldsymbol{x}), x_1 + \delta_1(\boldsymbol{x})] \times \cdots \times [x_d - \delta_d(\boldsymbol{x}), x_d + \delta_d(\boldsymbol{x})] \subset \overline{\Omega}$$

for all $\boldsymbol{x} = (x_1, \ldots, x_d) \in \overline{\Omega}$. The filter of v is defined by

$$\overline{v}(\boldsymbol{x}) = \frac{1}{a(\boldsymbol{x})} \int_{B(\boldsymbol{x})} \prod_{l=1}^{d} g_{\text{box}} \left(\frac{x_l}{2\delta_l(\boldsymbol{x})} \right) v(\boldsymbol{x} - \boldsymbol{z})\, d\boldsymbol{z}, \tag{8.53}$$

with

$$a(\boldsymbol{x}) = \prod_{l=1}^{d} 2\delta_l(\boldsymbol{x}).$$

Note that with this definition, the filter width in $\boldsymbol{x}$ in the direction x_l is $2\delta_l(\boldsymbol{x})$. □

Lemma 8.57 (Representation Formula for the Commutation Error) *Let $v \in C^1(\overline{U(\boldsymbol{x})})$, where $U(\boldsymbol{x})$ is a neighborhood of $\boldsymbol{x}$ such that $\boldsymbol{x} + B(\boldsymbol{x}) \subset U(\boldsymbol{x})$, and $\delta_l \in C^1(\overline{U(\boldsymbol{x})})$, $l = 1, \ldots, d$. Then, the i-th component of the commutation error has the form*

$$\left(\partial_i \overline{v} - \overline{\partial_i v}\right)(\boldsymbol{x}) = \sum_{l=1}^{d} \frac{\partial_i \delta_l(\boldsymbol{x})}{\delta_l(\boldsymbol{x})} \left(\overline{x_l \partial_l v} - x_l \overline{\partial_l v}\right)(\boldsymbol{x}). \tag{8.54}$$

Proof For the sake of simplifying the presentation, the proof will be given in one dimension. Because of the tensor product structure of the multi-dimensional filter, the proof for multiple dimensions uses the same techniques, see Berselli et al. (2007).

With the product rule of differentiation and the Leibniz rule, denoting with the prime the derivative, one gets

$$\begin{aligned}
\frac{d\overline{v}}{dx}(x) &= \frac{d}{dx}\left(\frac{1}{2\delta(x)}\int_{-\delta(x)}^{\delta(x)} v(x-z)\,dz\right) \\
&= -\frac{\delta'(x)}{2\delta^2(x)}\int_{-\delta(x)}^{\delta(x)} v(x-z)\,dz \\
&\quad + \frac{1}{2\delta(x)}\left[v(x-\delta(x))\delta'(x) - v(x+\delta(x))(-\delta'(x)\right] + \frac{1}{2\delta(x)}\int_{-\delta(x)}^{\delta(x)} v'(x-z)\,dz \\
&= -\frac{\delta'(x)}{2\delta^2(x)}\int_{-\delta(x)}^{\delta(x)} v(x-z)\,dz \\
&\quad + \frac{\delta'(x)}{2\delta(x)}\left[v(x-\delta(x)) + v(x+\delta(x))\right] + \overline{\frac{dv}{dx}}(x).
\end{aligned} \tag{8.55}$$

Since the bounds of the integral do not depend on z, integration by parts and the chain rule yields

$$\begin{aligned}
\int_{-\delta(x)}^{\delta(x)} zv'(x-z)\,dz &= -zv(x-z)\Big|_{z=-\delta(x)}^{z=\delta(x)} + \int_{-\delta(x)}^{\delta(x)} v(x-z)\,dz \\
&= -\delta(x)v(x-\delta(x)) - \delta(x)v(x+\delta(x)) + \int_{-\delta(x)}^{\delta(x)} v(x-z)\,dz.
\end{aligned}$$

Multiplying this identity by $\delta'(x)/(2\delta^2(x))$ and adding it to (8.55) gives

$$\begin{aligned}
&\left(\frac{d\overline{v}}{dx} - \overline{\frac{dv}{dx}}\right)(x) \\
&= -\frac{\delta'(x)}{2\delta^2(x)}\int_{-\delta(x)}^{\delta(x)} zv'(x-z)\,dz \qquad (8.56)\\
&= \frac{\delta'(x)}{\delta(x)}\left(\frac{1}{2\delta(x)}\int_{-\delta(x)}^{\delta(x)} (x-z)v'(x-z)\,dz - x\frac{1}{2\delta(x)}\int_{-\delta(x)}^{\delta(x)} v'(x-z)\,dz\right) \\
&= \frac{\delta'(x)}{\delta(x)}\left(\overline{xv'(x)} - x\,\overline{v'}\right).
\end{aligned}$$

This representation is exactly (8.54) for $d = 1$. ■

Theorem 8.58 (Pointwise Error Estimate of the Commutation Error) *Let* $v \in C^2(\overline{U(\boldsymbol{x})})$*, where* $U(\boldsymbol{x})$ *is defined in Lemma 8.57, and let* $\delta_l(\boldsymbol{x}) \in C^1(\overline{U(\boldsymbol{x})})$*,* $l = 1, \ldots, d$*, then*

$$\left|\partial_i \overline{v} - \overline{\partial_i v}\right|(\boldsymbol{x}) \le \frac{\|v\|_{C^2(\overline{U(\boldsymbol{x})})}}{3} \sum_{l=1}^{d} |\partial_i \delta_l(\boldsymbol{x})|\,|\delta_l(\boldsymbol{x})|. \tag{8.57}$$

Proof For the same reasons as in Lemma 8.57, the proof will be performed in one dimension. A Taylor series expansion with Lagrangian remainder gives

$$v'(x-z) = v'(x) - zv''(\xi), \quad \text{with some } \xi \in U(x).$$

Inserting this expression in (8.56) and using the definition of the norm in $C^2(\overline{U(x)})$, see Remark A.22, gives

$$\begin{aligned}
\left|\frac{d\overline{v}}{dx} - \overline{\frac{dv}{dx}}\right|(x) &= \left|\frac{\delta'(x)}{2\delta^2(x)} \int_{-\delta(x)}^{\delta(x)} zv'(x) - z^2 v''(\xi)\, dz\right| \\
&= \left|\frac{\delta'(x)}{2\delta^2(x)} \left(v'(x) \int_{-\delta(x)}^{\delta(x)} z\, dz - v''(\xi) \int_{-\delta(x)}^{\delta(x)} z^2\, dz\right)\right| \\
&\le \left|\frac{\delta'(x)}{2\delta^2(x)}\right| \|v\|_{C^2(\overline{U(x)})} \left|0 - \frac{2}{3}\delta^3(x)\right| \\
&= \frac{|\delta'(x)|\,\delta(x)}{3} \|v\|_{C^2(\overline{U(x)})}.
\end{aligned}$$

This estimate is the one-dimensional version of (8.57). ■

Remark 8.59 (Interpretation of Theorem 8.58) Estimate (8.57) shows that the commutation error vanishes if the filter width is constant in all directions, i.e., $\partial_i \delta_l(\boldsymbol{x}) = 0$, $i = 1, \ldots, d$. These conditions cannot be satisfied (or only trivially) if one considers a bounded domain and requires that the filter kernel should always be inside the closure of this domain. For simplicity, let $\Omega = (a, b) \subset \mathbb{R}$. If the filter kernel should be contained in $[a, b]$ then necessarily $\delta(x) \to 0$ as $x \to a$ and $\delta(x) \to 0$ as $x \to b$. Thus, either one has $\delta(x) = 0$, i.e., no filtering, or the filter width is not constant.

If the derivatives $\partial_i \delta_l(\boldsymbol{x}) = 0$, $i = 1, \ldots, d$, are bounded, then the commutation error tends to zero as $\delta_l(\boldsymbol{x}) \to 0$. □

Remark 8.60 (Further Results on Commutation Errors for Filters with Compact Support)

- In Berselli et al. (2007), the case of a general filter with compact support, variable filter width, and non-vanishing skewness is studied. A representation formula of the commutation error and a pointwise error estimate are derived.

- The analysis presented in this section requires a certain regularity of v, namely $v \in C^2(\overline{U(\boldsymbol{x})})$. In the case of turbulent flows, however, one cannot expect smooth solutions. In Berselli et al. (2007) also a pointwise estimate for the commutation error for functions with low regularity, concretely for Hölder-continuous functions, is proved.
- Since the filter width cannot be constant near the boundary, see Remark 8.59, a commutation error will be committed especially near the boundary. Estimates of this error for certain models of the mean velocity can be also found in Berselli et al. (2007). For the velocity, these models are called wall laws. An asymptotic analysis in Berselli and John (2006) shows for the turbulent channel flow, Example D.12, that the commutation error near the wall is at least as important as the divergence of the sgs stress tensor (8.29). Note that modeling the sgs stress tensor is the main issue in LES.

□

8.2.6 Summary of the Results Concerning Commutation Errors

Remark 8.61 (Summary of the Results Concerning the Commutation of Convolution and Differentiation) There are important situations in which the assumed commutation of filtering and differentiation is generally not true, e.g., if Ω is a bounded domain or if the filter width is not constant. Commuting these operators in these situation leads to extra terms. Omitting these terms, so-called commutation errors are committed.

Sections 8.2.3 and 8.2.4 considered the case of a bounded domain and the convolution with the Gaussian filter. First, the commutation error that is caused by extensions of functions from Ω to $\mathbb{R}^d$ which are not sufficiently smooth was derived in Sect. 8.2.3. In the following sections, an analysis of the commutation error in various norms for an arbitrary but fixed time was presented. In practical computations, the commutation error term is always neglected, expecting that it is small and vanishes if the filter width δ tends to zero. In Sect. 8.2.4, it was shown that this is not always true. The commutation error is asymptotically negligible in $L^p\left(\mathbb{R}^d\right)$, i.e., it vanishes as the averaging radius $\delta \to 0$, if and only if the normal stress vanishes almost everywhere on the boundary. In other words, it is asymptotically negligible in $L^p\left(\mathbb{R}^d\right)$ if and only if the fluid and the boundary exert zero force on each other. The expected convergence of the commutation error as $\delta \to 0$ was shown in the $H^{-1}(\Omega)$ norm and for a weak form of the commutation error.

When using the box filter, one can avoid the extension of functions off the domain. However, the required non-constant filter width causes also a commutation error. This error was analyzed in Sect. 8.2.5. It was shown that it depends on the variation of the filter width. □

Remark 8.62 (Practical Simulations) In practice, the commutation error is not of importance. By experience, it is known that LES models do not behave appropriately near boundaries. Thus, they are either modified, like by using the van Driest damping in the Smagorinsky model, see Remark 8.127, or even completely different approaches are used near the boundary. An overview on such approaches is given in Piomelli and Balaras (2002). Two possibilities are

- imposing some form of wall law,
- solving numerically a set of simplified equations in the boundary layer region, which is the so-called zonal approach.

□

8.3 Large Eddy Simulation: The Smagorinsky Model

Remark 8.63 (Motivation and Contents of this Section) The Smagorinsky model, proposed in Smagorinsky (1963), is one of the most popular turbulence models. From the mathematical point of view, the continuous Smagorinsky model is well understood. Existence and uniqueness of a solution in two and three dimensions can be proved, see Sect. 8.3.2. These results were obtained in Ladyženskaja (1967). Therefore, in the mathematical literature, the continuous model is often called Ladyzhenskaya model. In addition, a finite element error analysis can be performed, see Sect. 8.3.3. From the computational point of view, the Smagorinsky model is easy to implement, it has low computational cost, and generally it is quite robust in simulations. The last point means that the simulations do not blow up. However, as it shall be discussed in some detail in Sect. 8.3.4, the Smagorinsky model possesses a number of drawbacks. In particular, the computational results depend on a coefficient, the so-called Smagorinsky coefficient, whose good choice is in general situations an open problem.

In addition, it is part of some other turbulence models, like the three-scale coarse space projection-based variational multiscale model presented in Sect. 8.8.6. □

8.3.1 *The Model of the SGS Stress Tensor: Eddy Viscosity Models*

Remark 8.64 (Some Properties of the SGS Stress Tensor) The space-averaged Navier–Stokes equations are not yet closed and the divergence of the subgrid-scale stress tensor

$$\nabla \cdot \mathbb{T} = \nabla \cdot \left(\overline{\boldsymbol{u}\boldsymbol{u}^T} \right) - \nabla \cdot \left(\overline{\boldsymbol{u}}\, \overline{\boldsymbol{u}}^T \right) \tag{8.58}$$

needs to be modeled, see Remark 8.32. Models should possess some important properties of $\mathbb{T}$, as the followings.

- The subgrid stress tensor is symmetric, since $\boldsymbol{u}\boldsymbol{u}^T$ and $\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T$ are symmetric.
- The subgrid stress tensor is Galilean invariant. Consider two coordinate systems

  $$(t, \boldsymbol{x}) \quad \text{and} \quad (\hat{t}, \hat{\boldsymbol{x}}) = (t, \boldsymbol{x} - \boldsymbol{U}t), \tag{8.59}$$

  where $\boldsymbol{U}$ is a constant velocity. These systems describe two so-called inertial frames of references. The second frame is moving with constant velocity with respect to the first frame. Galilean invariance means that one gets the same expressions in both frames.
  Since $\boldsymbol{u}$ is the velocity (derivative with respect to time) of $\boldsymbol{x}$, it follows from (8.59) by differentiating $\hat{\boldsymbol{x}}$ with respect to time that

  $$\hat{\boldsymbol{u}}(\hat{t}, \hat{\boldsymbol{x}}) = \boldsymbol{u}(t, \boldsymbol{x}) - \boldsymbol{U}. \tag{8.60}$$

  In the special case $\boldsymbol{u} = \boldsymbol{U}$, the coordinate system $(\hat{t}, \hat{\boldsymbol{x}})$ moves with the flow. This case is the so-called Lagrangian point of view of a flow field, i.e., one observes the motion of the fluid such that one follows an individual "fluid particle" as it moves through time and space, see also Remark 8.190.
  Using the chain rule, (8.60), and (8.59), one obtains for the partial derivative of the velocity

  $$\frac{\partial \hat{\boldsymbol{u}}_j}{\partial \hat{x}_i} = \frac{\partial \boldsymbol{u}_j}{\partial t}\frac{\partial t}{\partial \hat{x}_i} + \sum_{k=1}^{d} \frac{\partial \boldsymbol{u}_j}{\partial x_k}\frac{\partial x_k}{\partial \hat{x}_i} = \frac{\partial \boldsymbol{u}_j}{\partial x_i}, \quad i, j = 1, \ldots, d.$$

  Since the sgs stress tensor is composed only of partial derivative of the velocity, it is Galilean invariant.
  With similar calculations, it can be shown that the Navier–Stokes equations (8.1) are Galilean invariant (with $\hat{\boldsymbol{u}}$ instead of $\boldsymbol{u}$).

Further requirements for satisfactory turbulence models are discussed in Berselli et al. (2006, Sect. 6.2). □

Remark 8.65 (Modeling of the Deviatoric Part) For incompressible fluids, the pressure is the trace of the stress tensor multiplied with $-1/3$, see (2.19), and the velocity part of the stress tensor is trace-free. Similarly, one considers only the deviatoric or trace-free sgs stress tensor

$$\mathbb{T} - \frac{\mathbb{T}_{11} + \mathbb{T}_{22} + \mathbb{T}_{33}}{3} \tag{8.61}$$

and the second term is usually added to the filtered pressure $\overline{p}$

$$\overline{p} + \frac{\mathbb{T}_{11} + \mathbb{T}_{22} + \mathbb{T}_{33}}{3}.$$

Hence, it remains to model the deviatoric part (8.61) of the sgs stress tensor.

For simplicity of notation, the modified pressure will be also denoted by $\overline{p}$. □

Remark 8.66 (The Boussinesq Hypothesis) The starting point of the derivation of the Smagorinsky model is the Boussinesq hypothesis stated in Boussinesq (1877): "Turbulent fluctuations are dissipative in the mean". This hypothesis is based upon the resemblance of the elastic collisions of molecules with the interaction of small scales in flows. Expressing this hypothesis for the model of the sgs stress tensor, it has the form

$$\mathbb{T} - \frac{\mathbb{T}_{11} + \mathbb{T}_{22} + \mathbb{T}_{33}}{3} = -\nu_{\mathrm{T}} \mathbb{D}\left(\overline{\boldsymbol{u}}\right), \tag{8.62}$$

where $\nu_{\mathrm{T}} \geq 0$ is called turbulent viscosity or eddy viscosity. From (8.62), one can see that this model introduces a viscous term. Usually, the turbulent viscosity depends on the solution, hence the new viscous term is nonlinear. The term on the right-hand side of (8.62) is often called eddy viscosity model. □

Remark 8.67 (Modeling the Turbulent Viscosity ν_{T}) It is known for the dissipation of turbulent energy that $\varepsilon \sim U^3/L$, see (8.2). For the Kolmogorov scales, one obtains from (8.5) with a direct calculation that $\varepsilon = u_\lambda^3/\lambda$. It is now assumed that not only for the largest and smallest scales a relation of this type holds but for every length scale and the corresponding velocity scale. In particular, it is assumed that

$$\varepsilon \sim \frac{U_{\mathrm{int}}^3}{L_{\mathrm{int}}}, \quad \varepsilon \sim \frac{U_\delta^3}{\delta}, \tag{8.63}$$

where δ is the filter width. The scale L_{int} is the so-called integral length scale, which characterizes the distance over which the small scales are correlated, and U_{int} is the corresponding velocity scale. One obtains directly from (8.63) the relation

$$U_\delta \sim U_{\mathrm{int}} \left(\frac{\delta}{L_{\mathrm{int}}}\right)^{1/3}. \tag{8.64}$$

The goal of the turbulence model is to capture scales of size δ. The Reynolds number of these scales should be 1, i.e., it should hold

$$\mathrm{Re}(\delta) = \frac{\delta U_\delta}{\nu_{\mathrm{T}}} = 1.$$

Hence, one gets with (8.64)

$$\nu_{\mathrm{T}} = U_\delta \delta \sim U_{\mathrm{int}} L_{\mathrm{int}}^{-1/3} \delta^{4/3}. \tag{8.65}$$

The next assumption correlates the integral velocity and length scales. It is assumed that the integral velocity scale depends linearly on the norm of the deformation tensor of the filtered velocity

$$U_{\text{int}} \sim L_{\text{int}} \left\| \mathbb{D}\left(\overline{\boldsymbol{u}}\right)\right\|_{\text{F}}.$$

Inserting this assumption in (8.65) and replacing similarity by an unknown coefficient yields

$$\nu_{\text{T}} = C L_{\text{int}}^{2/3} \delta^{4/3} \left\| \mathbb{D}\left(\overline{\boldsymbol{u}}\right)\right\|_{\text{F}}.$$

The integral length scale L_{int} is usually hard to determine. Thus, one uses the approximation $L_{\text{int}} \sim \delta$ to get, with a modified constant,

$$\nu_{\text{T}} = C_S \delta^2 \left\| \mathbb{D}\left(\overline{\boldsymbol{u}}\right)\right\|_{\text{F}}. \tag{8.66}$$

□

Remark 8.68 (The Smagorinsky Model) With (8.66), the Smagorinsky model introduces the following additional term to the left-hand side of the momentum equation of the Navier–Stokes equations

$$-\nabla \cdot \left(C_S \delta^2 \left\| \mathbb{D}\left(\overline{\boldsymbol{u}}\right)\right\|_{\text{F}} \mathbb{D}\left(\overline{\boldsymbol{u}}\right)\right), \tag{8.67}$$

where $C_S \geq 0$ is the dimensionless Smagorinsky coefficient and δ is related to the (local) mesh width. In particular in engineering literature, it is common to write the Smagorinsky model in the form

$$-\nabla \cdot \left(2 \left(C_S^* \delta \right)^2 \left(2\mathbb{D}\left(\overline{\boldsymbol{u}}\right) : \mathbb{D}\left(\overline{\boldsymbol{u}}\right)\right)^{1/2} \mathbb{D}\left(\overline{\boldsymbol{u}}\right)\right). \tag{8.68}$$

Since the Smagorinsky model contains only the deformation tensor of the velocity, it is a symmetric model. It is also Galilean invariant, since only first order spatial derivatives of the velocity appear, compare Remark 8.64.

Replacing the space-averaged velocity and pressure $(\overline{\boldsymbol{u}}, \overline{p})$ with their approximations $(\boldsymbol{w}, r)$, the momentum balance of the Smagorinsky model has the form

$$\partial_t \boldsymbol{w} - \nabla \cdot \left(\left(\nu + C_S \delta^2 \left\| \mathbb{D}\left(\boldsymbol{w}\right)\right\|_{\text{F}} \right) \mathbb{D}\left(\boldsymbol{w}\right)\right) + \left(\boldsymbol{w} \cdot \nabla\right) \boldsymbol{w} + \nabla r = \boldsymbol{f} \tag{8.69}$$

in $(0, T] \times \Omega$. □

Remark 8.69 (The Smagorinsky Filter) The use of the Smagorinsky model does not require the specification of a concrete filter. The filter does not appear explicitly, only the filter width. However, it can be shown for homogeneous isotropic turbulence that there is a uniquely implied filter, see Pope (2000, Sect. 13.4.3). □

8.3.2 Existence and Uniqueness of a Solution of the Continuous Smagorinsky Model

Remark 8.70 (To the Proof of the Existence and Uniqueness of a Solution) The proof of the existence and uniqueness of a weak solution of the continuous Smagorinsky model was given in Ladyženskaja (1967), see also Ladyzhenskaya (1969, Supplement 1) for an overview. The presentation in this section follows also John (2004, Sect. 6.1).

The existence of a weak solution is proved with the Galerkin method, which was used in Sect. 7.1 to show the existence of a weak solution of the Navier–Stokes equations. Because of the nonlinear viscous term of the Smagorinsky model, one needs different spaces than for the Navier–Stokes equations to ensure that all terms are well defined. In fact, the velocity space needed for the Smagorinsky model is of somewhat higher regularity than the space (7.5) used for the Navier–Stokes equations. This additional regularity simplifies the technical tools needed in the proof, in particular the convergence of the nonlinear convective term, compare Lemma 8.83 with Lemmas 7.15 and 7.16. The main issue consists in proving the convergence of the nonlinear viscous term of the Smagorinsky model, see the considerations from Lemmas 8.84 to 8.90. It can be shown that the Smagorinsky model defines a monotone operator and the theory of such operators can be applied. Even more important, the higher regularity allows to prove the uniqueness of the weak solution.

Altogether, the main questions concerning the analysis of the Smagorinsky model are answered, in contrast to the Navier–Stokes equations. This situation indicates that the Smagorinsky model is in some sense an easier problem than the Navier–Stokes equations, despite of the additional nonlinear term. In this respect, it behaves exactly the way a turbulence model should do, compare Remark 8.22. □

Remark 8.71 (Notation) The notation in this section will indicate that the solution of the Smagorinsky model is generally not the solution of the Navier–Stokes equations. To this end, the velocity field will be denoted by $\boldsymbol{w}$ and the pressure by r. □

Remark 8.72 (The Strong Form of the Smagorinsky Model) Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be a bounded domain with Lipschitz boundary Γ. For the sake of simplifying the analysis a little bit, the gradient form of the Smagorinsky model is considered. This model, equipped with homogeneous Dirichlet boundary conditions, has the form

$$
\begin{aligned}
\partial_t \boldsymbol{w} - \nabla \cdot \left(\left(\nu + \nu_{\mathrm{Sm}} \left\| \nabla \boldsymbol{w} \right\|_{\mathrm{F}} \right) \nabla \boldsymbol{w} \right) + \left(\boldsymbol{w} \cdot \nabla \right) \boldsymbol{w} + \nabla r &= \boldsymbol{f} \quad \text{in } (0, T] \times \Omega, \\
\nabla \cdot \boldsymbol{w} &= 0 \quad \text{in } (0, T] \times \Omega, \\
\boldsymbol{w} &= \boldsymbol{0} \quad \text{in } (0, T] \times \Gamma, \\
\boldsymbol{w}(0, \cdot) &= \boldsymbol{w}_0 \ \text{in } \Omega, \\
\int_\Omega r \, d\boldsymbol{x} &= 0 \quad \text{in } (0, T],
\end{aligned}
\tag{8.70}
$$

with $\nu_{\mathrm{Sm}} \in \mathbb{R}$, $\nu_{\mathrm{Sm}} > 0$, $\boldsymbol{f} \in L^2\left(0,T;L^2(\Omega)\right)$, given initial condition $\boldsymbol{w}_0$, and finite final time $T < \infty$. Because a bounded domain with Lipschitz boundary is considered, the Sobolev imbedding of $W^{1,3}(\Omega) \to L^q(\Omega)$, $q \in [1,\infty)$, compare (A.16) is compact, see Theorem A.42 vii). □

Remark 8.73 (Function Spaces for the Weak Formulation) For achieving that all terms in the weak formulation are well defined, the Banach space

$$W^{1,3}_{0,\mathrm{div}}(\Omega) = \{\boldsymbol{v} \in W^{1,3}(\Omega) \; : \; \boldsymbol{v}|_{\partial\Omega} = \boldsymbol{0}\,, \; \nabla\cdot\boldsymbol{v} = 0 \text{ in } \Omega\} \tag{8.71}$$

is used. This space is equipped with the same norm as $W^{1,3}_0(\Omega)$, see Remark A.31. The appropriate velocity space is

$$V = L^3\left(0,T;W^{1,3}_{0,\mathrm{div}}(\Omega)\right) \cap H^1\left(0,T;L^2(\Omega)\right) \tag{8.72}$$

equipped with the norm

$$\|\boldsymbol{v}\|_V = \|\nabla\boldsymbol{v}\|_{L^3(0,T;L^3(\Omega))} + \|\partial_t\boldsymbol{v}\|_{L^2(0,T;L^2(\Omega))}\,.$$

Compared with the spaces used for the Navier–Stokes equations, see (7.5), one can observe the higher regularity of the spaces used in the analysis of the Smagorinsky model. □

Remark 8.74 (Weak Formulation) The weak formulation of the Smagorinsky model is obtained in the usual way. The strong formulation (8.70) is multiplied with appropriate test functions, the resulting equations are integrated in Ω, and integration by parts is applied to transfer derivatives from the solution to the test functions. As for the Navier–Stokes equations, the test functions are divergence-free such that the pressure term vanishes.

The weak formulation of the Smagorinsky model reads as follows: Find $\boldsymbol{w} \in V$ such that $\boldsymbol{w}(0,\boldsymbol{x}) = \boldsymbol{w}_0 \in W^{1,3}_{0,\mathrm{div}}(\Omega)$ and for all $\boldsymbol{v} \in V$

$$\int_0^T (\partial_t\boldsymbol{w} + (\boldsymbol{w}\cdot\nabla)\,\boldsymbol{w}, \boldsymbol{v}) + ((\nu + \nu_{\mathrm{Sm}}\,\|\nabla\boldsymbol{w}\|_{\mathrm{F}})\,\nabla\boldsymbol{w}, \nabla\boldsymbol{v}) \; dt = \int_0^T (\boldsymbol{f}, \boldsymbol{v}) \; dt. \tag{8.73}$$

The function $\boldsymbol{w} \in V$ is called weak solution of the Smagorinsky model. □

Remark 8.75 (Preparations for the Convergence Proof) The existence proof is based on three a priori or stability error estimates that will be given in Lemmas 8.76–8.78. □

Lemma 8.76 (Stability Estimate for $\|\boldsymbol{w}\|_{L^\infty(0,T;L^2(\Omega))}$) *Let $(\boldsymbol{w}, r) \in V \times Q$, with $Q = L^2_0(\Omega)$, $\boldsymbol{f} \in L^2\left(0,T;L^2(\Omega)\right)$, and $\boldsymbol{w}_0 \in W^{1,3}_{0,\mathrm{div}}(\Omega)$. Then, each solution of (8.70) satisfies*

$$\|\boldsymbol{w}(T)\|_{L^2(\Omega)} \le \|\boldsymbol{w}_0\|_{L^2(\Omega)} + \|\boldsymbol{f}\|_{L^1(0,T;L^2(\Omega))} \qquad \forall\; T > 0. \tag{8.74}$$

Proof Since $(\boldsymbol{w}, r)$ is a solution of (8.70), one obtains by testing the momentum equation of (8.70) with $\boldsymbol{w}$ and integration by parts

$$(\partial_t \boldsymbol{w}, \boldsymbol{w}) + ((\nu + \nu_{\mathrm{Sm}} \|\nabla \boldsymbol{w}\|_{\mathrm{F}}) \nabla \boldsymbol{w}, \nabla \boldsymbol{w}) + n(\boldsymbol{w}, \boldsymbol{w}, \boldsymbol{w}) - (r, \nabla \cdot \boldsymbol{w}) = (\boldsymbol{f}, \boldsymbol{w}).$$

Because $\nabla \cdot \boldsymbol{w} = 0$, the convective term and the term with the pressure r vanish, such that

$$\frac{1}{2}\frac{d}{dt}(\boldsymbol{w}, \boldsymbol{w}) + ((\nu + \nu_{\mathrm{Sm}} \|\nabla \boldsymbol{w}\|_{\mathrm{F}}) \nabla \boldsymbol{w}, \nabla \boldsymbol{w}) = (\boldsymbol{f}, \boldsymbol{w}). \tag{8.75}$$

Since $0 < \nu + \nu_{\mathrm{Sm}} \|\nabla \boldsymbol{w}\|_{\mathrm{F}}$, the second term on the right-hand side is non-negative. Neglecting this term, applying the chain rule for the first term of the left-hand side and the Cauchy–Schwarz inequality (A.10) for the right-hand side gives

$$\|\boldsymbol{w}\|_{L^2(\Omega)} \frac{d}{dt} \|\boldsymbol{w}\|_{L^2(\Omega)} \le \|\boldsymbol{f}\|_{L^2(\Omega)} \|\boldsymbol{w}\|_{L^2(\Omega)}.$$

Cancellation of $\|\boldsymbol{w}\|_{L^2(\Omega)}$, integration on $(0, T)$, and noting that for a finite time interval $\boldsymbol{f} \in L^2\left(0, T; L^2(\Omega)\right)$ implies $\boldsymbol{f} \in L^1\left(0, T; L^2(\Omega)\right)$ completes the proof. ∎

Lemma 8.77 (Stability of $\|\nabla \boldsymbol{w}\|_{L^3\left(0,T;L^3(\Omega)\right)}$) *With the same assumptions as in Lemma 8.76, it holds for all $T > 0$*

$$\begin{aligned} &\|\boldsymbol{w}(T)\|_{L^2(\Omega)}^2 + 2\int_0^T ((\nu + \nu_{\mathrm{Sm}} \|\nabla \boldsymbol{w}\|_{\mathrm{F}}) \nabla \boldsymbol{w}, \nabla \boldsymbol{w})\ dt \\ &\le 2\|\boldsymbol{w}_0\|_{L^2(\Omega)}^2 + 3\|\boldsymbol{f}\|_{L^1\left(0,T;L^2(\Omega)\right)}^2 = C_1(T). \end{aligned} \tag{8.76}$$

In particular, it is $\|\nabla \boldsymbol{w}\|_{L^3\left(0,T;L^3(\Omega)\right)} \le \tilde{C}_1(T)$.

Proof Starting with (8.75), one gets by integration on $(0, T)$

$$\|\boldsymbol{w}(T)\|_{L^2(\Omega)}^2 + 2\int_0^T ((\nu + \nu_{\mathrm{Sm}} \|\nabla \boldsymbol{w}\|_{\mathrm{F}}) \nabla \boldsymbol{w}, \nabla \boldsymbol{w})\ dt = \|\boldsymbol{w}_0\|_{L^2(\Omega)}^2 + 2\int_0^T (\boldsymbol{f}, \boldsymbol{w})\ dt, \tag{8.77}$$

which is an energy inequality like (7.27), here even an energy equality. Applying the Cauchy–Schwarz inequality (A.10) and inequality (8.74), which is valid for all times t, it follows for the second term on the right-hand side that

$$\begin{aligned} &\int_0^T (\boldsymbol{f}, \boldsymbol{w})\ dt \\ &\le \int_0^T \|\boldsymbol{w}(t)\|_{L^2(\Omega)} \|\boldsymbol{f}(t)\|_{L^2(\Omega)}\ dt \\ &\le \|\boldsymbol{w}_0\|_{L^2(\Omega)} \int_0^T \|\boldsymbol{f}(t)\|_{L^2(\Omega)}\ dt + \int_0^T \|\boldsymbol{f}(t)\|_{L^2(\Omega)} \left(\int_0^t \|\boldsymbol{f}(t')\|_{L^2(\Omega)}\ dt'\right) dt. \end{aligned}$$

Using Young's inequality (A.5) and the non-negativeness of $\|\boldsymbol{f}(t')\|_{L^2(\Omega)}$ yields

$$\begin{aligned}\int_0^T (\boldsymbol{f}(t,\boldsymbol{x}), \boldsymbol{w}(t,\boldsymbol{x}))\, dt &\le \frac{\|\boldsymbol{w}_0\|_{L^2(\Omega)}^2}{2} + \frac{1}{2}\left(\int_0^T \|\boldsymbol{f}(t)\|_{L^2(\Omega)}\, dt\right)^2 \\ &\quad + \int_0^T \|\boldsymbol{f}(t)\|_{L^2(\Omega)} \left(\int_0^T \|\boldsymbol{f}(t')\|_{L^2(\Omega)}\, dt'\right) dt \\ &= \frac{\|\boldsymbol{w}_0\|_{L^2(\Omega)}^2}{2} + \frac{3}{2}\|\boldsymbol{f}\|_{L^1(0,T;L^2)}^2 .\end{aligned}$$

Inserting this estimate in (8.77) gives (8.76).

It follows that

$$\begin{aligned}2\nu_{\mathrm{Sm}}\int_0^T (\|\nabla\boldsymbol{w}\|_{\mathrm{F}}\nabla\boldsymbol{w}, \nabla\boldsymbol{w})\, dt &= 2\nu_{\mathrm{Sm}}\int_0^T \|\nabla\boldsymbol{w}\|_{\mathrm{F}}^3\, d\boldsymbol{x} \\ &= 2\nu_{\mathrm{Sm}}\|\nabla\boldsymbol{w}\|_{L^3(0,t;L^3(\Omega))}^3 \le C_1(T),\end{aligned}$$

which proves the second statement of the lemma. ■

Lemma 8.78 (Stability of $\|\boldsymbol{w}\|_{H^1(0,T;L^2(\Omega))}$) *If the assumptions of Lemma 8.76 are valid, then*

$$\|\nabla\boldsymbol{w}(T)\|_{L^3(\Omega)}^3 + \frac{3}{2\nu_{\mathrm{Sm}}}\|\boldsymbol{w}\|_{H^1(0,T;L^2(\Omega))}^2 \le C_2(T). \tag{8.78}$$

In particular, it follows that $\|\boldsymbol{w}\|_{H^1(0,T;L^2(\Omega))} \le \tilde{C}_2(T)$.

Proof To prove (8.78), one starts by testing the momentum equation of (8.70) with $\partial_t\boldsymbol{w}$ and integrating in $(0,T)$. With the chain rule, one gets

$$\frac{1}{3}\int_\Omega \frac{d}{dt}\|\nabla\boldsymbol{w}\|_{\mathrm{F}}^3\, d\boldsymbol{x} = (\|\nabla\boldsymbol{w}\|_{\mathrm{F}}\nabla\boldsymbol{w}, \partial_t\nabla\boldsymbol{w}) = (\|\nabla\boldsymbol{w}\|_{\mathrm{F}}\nabla\boldsymbol{w}, \nabla\partial_t\boldsymbol{w}).$$

This relation, integration by parts, using that $\partial_t\boldsymbol{w}$ is weakly divergence-free, and applying formulas of type (7.13) yields

$$\begin{aligned}&\int_0^T \|\partial_t\boldsymbol{w}\|_{L^2(\Omega)}^2\, dt + \frac{\nu}{2}(\nabla\boldsymbol{w}(T), \nabla\boldsymbol{w}(T)) + \frac{\nu_{\mathrm{Sm}}}{3}\int_\Omega \|\nabla\boldsymbol{w}(T)\|_{\mathrm{F}}^3\, d\boldsymbol{x} \\ &= \frac{\nu}{2}(\nabla\boldsymbol{w}_0, \nabla\boldsymbol{w}_0) + \frac{\nu_{\mathrm{Sm}}}{3}\int_\Omega \|\nabla\boldsymbol{w}_0\|_{\mathrm{F}}^3\, d\boldsymbol{x} - \int_0^T n(\boldsymbol{w},\boldsymbol{w},\partial_t\boldsymbol{w})\, dt + \int_0^T (\boldsymbol{f}, \partial_t\boldsymbol{w})\, dt.\end{aligned} \tag{8.79}$$

First, the convective term will be estimated. Hölder's inequality (A.9) gives

$$\begin{aligned}\int_\Omega \left(\boldsymbol{w}^T\boldsymbol{w}\right)(\nabla\boldsymbol{w} : \nabla\boldsymbol{w})\ d\boldsymbol{x} &= \int_\Omega \boldsymbol{w}^2 (\nabla\boldsymbol{w})^2\ d\boldsymbol{x} \\ &\leq \left\|\boldsymbol{w}^2\right\|_{L^3(\Omega)} \left\|(\nabla\boldsymbol{w})^2\right\|_{L^{3/2}} \\ &= \|\boldsymbol{w}\|_{L^6(\Omega)}^2 \|\nabla\boldsymbol{w}\|_{L^3(\Omega)}^2 .\end{aligned}$$

With the Sobolev imbedding $W^{1,3}(\Omega) \to L^6(\Omega)$, see (A.16), and Poincaré's inequality (A.12), it follows that

$$\int_\Omega \left(\boldsymbol{w}^T\boldsymbol{w}\right)(\nabla\boldsymbol{w} : \nabla\boldsymbol{w})\ d\boldsymbol{x} \leq C \|\nabla\boldsymbol{w}\|_{L^3(\Omega)}^4 . \tag{8.80}$$

Applying Young's inequality, one gets

$$\begin{aligned}&\left(w_1\partial_x w_1 + w_2\partial_y w_1 + w_3\partial_z w_1\right)^2 \\ &= w_1^2(\partial_x w_1)^2 + w_2^2\left(\partial_y w_1\right)^2 + w_3^2(\partial_z w_1)^2 + 2w_1w_2\partial_x w_1\partial_y w_1 + 2w_1w_3\partial_x w_1\partial_z w_1 \\ &\quad + 2w_2w_3\partial_y w_1\partial_z w_1 \\ &\leq w_1^2(\partial_x w_1)^2 + w_2^2\left(\partial_y w_1\right)^2 + w_3^2(\partial_z w_1)^2 + w_1^2\left(\partial_y w_1\right)^2 + w_2^2(\partial_x w_1)^2 + w_1^2(\partial_z w_1)^2 \\ &\quad + w_3^2(\partial_x w_1)^2 + w_2^2(\partial_z w_1)^2 + w_3^2\left(\partial_y w_1\right)^2 \\ &= \left(w_1^2 + w_2^2 + w_3^2\right)\left((\partial_x w_1)^2 + \left(\partial_y w_1\right)^2 + (\partial_x w_1)^2\right).\end{aligned}$$

Using now once more Young's inequality and inserting estimate (8.80) leads to

$$\begin{aligned}\int_0^T n(\boldsymbol{w}, \boldsymbol{w}, \partial_t\boldsymbol{w})\ dt &= \int_0^T \int_\Omega (\boldsymbol{w}\cdot\nabla)\,\boldsymbol{w}\cdot\partial_t\boldsymbol{w}\ d\boldsymbol{x}\ dt \\ &\leq \int_0^T \int_\Omega \left(\frac{(\partial_t\boldsymbol{w})^2}{4} + ((\boldsymbol{w}\cdot\nabla)\,\boldsymbol{w})^2\right) d\boldsymbol{x}\ dt \\ &\leq \int_0^T \int_\Omega \left(\frac{\partial_t\boldsymbol{w}^2}{4} + \left(\boldsymbol{w}^T\boldsymbol{w}\right)(\nabla\boldsymbol{w} : \nabla\boldsymbol{w})\right) d\boldsymbol{x}\ dt \\ &\leq \frac{1}{4}\|\partial_t\boldsymbol{w}\|_{L^2(0,T;L^2(\Omega))}^2\ dt + C\int_0^T \|\nabla\boldsymbol{w}\|_{L^3(\Omega)}^4\ dt.\end{aligned}$$

Young's inequality gives also

$$\begin{aligned}\int_0^T (\boldsymbol{f}, \partial_t\boldsymbol{w})\ dt &\leq \int_0^T \int_\Omega \left(\boldsymbol{f}^2 + \frac{(\partial_t\boldsymbol{w})^2}{4}\right) d\boldsymbol{x}\ dt \\ &= \|\boldsymbol{f}\|_{L^2(0,T;L^2(\Omega))}^2 + \frac{1}{4}\|\partial_t\boldsymbol{w}\|_{L^2(0,T;L^2(\Omega))}^2 .\end{aligned}$$

Inserting all estimates in (8.79) yields

$$\|\partial_t \boldsymbol{w}\|^2_{L^2(0,T;L^2(\Omega))} + \nu \|\nabla \boldsymbol{w}(T)\|^2_{L^2(\Omega)} + \frac{2\nu_{\mathrm{Sm}}}{3} \|\nabla \boldsymbol{w}(T)\|^3_{L^3(\Omega)} \tag{8.81}$$

$$\leq \nu \|\nabla \boldsymbol{w}_0\|^2_{L^2(\Omega)} + \frac{2\nu_{\mathrm{Sm}}}{3}\|\nabla \boldsymbol{w}_0\|^3_{L^3(\Omega)} + 2\|\boldsymbol{f}\|^2_{L^2(0,T;L^2(\Omega))} + 2C\int_0^T \|\nabla \boldsymbol{w}\|^4_{L^3(\Omega)}\ dt.$$

In particular, it follows that

$$\|\nabla \boldsymbol{w}(T)\|^3_{L^3(\Omega)} \leq \frac{3\nu}{2\nu_{\mathrm{Sm}}}\|\nabla \boldsymbol{w}_0\|^2_{L^2(\Omega)} + \|\nabla \boldsymbol{w}_0\|^3_{L^3(\Omega)} + \frac{3}{\nu_{\mathrm{Sm}}}\|\boldsymbol{f}\|^2_{L^2(0,T;L^2(\Omega))}$$
$$+\frac{3C}{\nu_{\mathrm{Sm}}}\int_0^T \|\nabla \boldsymbol{w}\|^4_{L^3(\Omega)}\ dt.$$

The application of Gronwall's lemma, Lemma A.53, gives

$$\|\nabla \boldsymbol{w}(T)\|^3_{L^3(\Omega)} \leq \left(\frac{3\nu}{2\nu_{\mathrm{Sm}}}\|\nabla \boldsymbol{w}_0\|^2_{L^2(\Omega)} + \|\nabla \boldsymbol{w}_0\|^3_{L^3(\Omega)} + \frac{3}{\nu_{\mathrm{Sm}}}\|\boldsymbol{f}\|^2_{L^2(0,T;L^2(\Omega))}\right)$$
$$\times \exp\left(\frac{3C}{\nu_{\mathrm{Sm}}}\int_0^T \|\nabla \boldsymbol{w}\|_{L^3(\Omega)}\ dt\right).$$

One still has to bound the term in the exponential. Using inequality (A.4) yields

$$\int_0^T \|\nabla \boldsymbol{w}\|_{L^3(\Omega)}\ dt = \int_0^T \left(\int_\Omega \sum_{i,j=1}^d \left|\frac{\partial w_i}{\partial x_j}\right|^3\ d\boldsymbol{x}\right)^{1/3}\ dt$$
$$\leq \int_0^T \left(\int_\Omega \left(\sum_{i,j=1}^d \left|\frac{\partial w_i}{\partial x_j}\right|^2\right)^{3/2}\ d\boldsymbol{x}\right)^{1/3}\ dt. \tag{8.82}$$

On the other hand, by the definition of the Frobenius norm, it is

$$\int_\Omega \|\nabla \boldsymbol{w}\|_{\mathrm{F}}(\nabla \boldsymbol{w} : \nabla \boldsymbol{w})\ d\boldsymbol{x} = \int_\Omega (\nabla \boldsymbol{w} : \nabla \boldsymbol{w})^{3/2}\ d\boldsymbol{x} = \int_\Omega \left(\sum_{i,j=1}^d \left|\frac{\partial w_i}{\partial x_j}\right|^2\right)^{3/2}\ d\boldsymbol{x}.$$

Inserting this identity in (8.82), one obtains with Hölder's inequality (A.9) and the a priori estimate (8.76)

$$\int_0^T \|\nabla \boldsymbol{w}\|_{L^3(\Omega)}\ dt \leq \int_0^T \left(\int_\Omega \|\nabla \boldsymbol{w}\|_{\mathrm{F}}(\nabla \boldsymbol{w} : \nabla \boldsymbol{w})\ d\boldsymbol{x}\right)^{1/3}\ dt$$
$$\leq \left(\int_0^T \int_\Omega \|\nabla \boldsymbol{w}\|_{\mathrm{F}}(\nabla \boldsymbol{w} : \nabla \boldsymbol{w})\ d\boldsymbol{x}\, dt\right)^{1/3}\left(\int_0^T dt\right)^{2/3}$$
$$\leq T^{2/3}\left(\frac{C_1(T)}{2\nu_{\mathrm{Sm}}}\right)^{1/3},$$

which gives

$$\|\nabla \boldsymbol{w}(T)\|^3_{L^3(\Omega)} \le C(T).$$

Together with (8.81), estimate (8.78) follows. ■

Remark 8.79 (Setup of the Finite-Dimensional Problem) A sequence $\{\boldsymbol{w}^n\} \subset V$ will be constructed, where the $\boldsymbol{w}^n$ is the unique solution of a Smagorinsky problem in a space with dimension n. It will be shown that a subsequence converges to a solution $\boldsymbol{w} \in V$ of (8.73).

Let $\{\boldsymbol{v}^n_l(\boldsymbol{x})\}_{l=1}^{\infty} \subset W^{1,3}_{0,\mathrm{div}}$ be a sequence of linearly independent functions that are orthonormal with respect to the $L^2(\Omega)$ inner product in Ω and with $\boldsymbol{v}^1_l = \boldsymbol{w}_0$, $l = 1, \ldots, \infty$. Then, the solution of the Smagorinsky problem in the space spanned by $\{\boldsymbol{v}^n_l(\boldsymbol{x})\}_{l=1}^{n}$ is sought in the form

$$\boldsymbol{w}^n(t,\boldsymbol{x}) = \sum_{l=1}^{n} \alpha^n_l(t)\,\boldsymbol{v}^n_l(\boldsymbol{x}) \tag{8.83}$$

satisfying

$$\alpha^n_l(0) = \begin{cases} 1, & l = 1, \\ 0, & l > 1, \end{cases}$$

and

$$\left(\partial_t \boldsymbol{w}^n, \boldsymbol{v}^n_l\right) + \left(\left(\nu + \nu_{\mathrm{Sm}}\|\nabla \boldsymbol{w}^n\|_{\mathrm{F}}\right)\nabla \boldsymbol{w}^n, \nabla \boldsymbol{v}^n_l\right) + n\left(\boldsymbol{w}^n, \boldsymbol{w}^n, \boldsymbol{v}^n_l\right) = \left(\boldsymbol{f}, \boldsymbol{v}^n_l\right), \tag{8.84}$$

$l = 1, \ldots, n$. System (8.84) is an autonomous, quasi-linear system of ordinary differential equations with respect to the unknown functions $\alpha^n_l(t)$. □

Lemma 8.80 (Existence and Uniqueness of a Solution of the Problem in the Finite-Dimensional Space) *System (8.84) admits a unique solution for all $T > 0$. Moreover, the estimates (8.74), (8.76), and (8.78) are valid for $\boldsymbol{w}^n$, where the right-hand sides do not depend on n. Hence $\boldsymbol{w}^n \in V$.*

Proof Like for the Navier–Stokes equations, the proof is based on the theorem of Carathéodory, see Theorem A.50. The first part can be taken literally from the proof of Lemma 7.10.

One has to show a Lipschitz condition for the right-hand side of

$$\frac{d\alpha^n_l}{dt}(t) = F\left(\alpha^n_l\right), \quad t \in (0,T], \tag{8.85}$$

with $F \in L^2(0,T)$. The functions α^n_l appear linearly and quadratically on the right-hand side of (8.85). Hence, the Lipschitz condition is satisfied, since linear and quadratic functions are Lipschitz continuous. Consequently, the local existence and uniqueness of an absolutely continuous solution $\boldsymbol{w}^n(t,\boldsymbol{x})$ in some maximal interval

$[0, t_n]$ with $0 < t_n \le T$ can be concluded from the theorem of Carathéodory. If $t_n < T$, then $\boldsymbol{w}^n(t)$ blows up as $t \to t_n$.

Now, stability estimates have be proved which guarantee that this situation cannot happen and therefore $t_n = T$. The boundedness of

$$\sup_{[0,T]} \sum_{l=1}^{n} \left(\alpha_l^n\right)^2 (t)$$

has to be shown. From the $L^2(\Omega)$ orthonormality of $\{\boldsymbol{v}_l^n(\boldsymbol{x})\}$, it follows that

$$\sup_{[0,T]} \sum_{l=1}^{n} \left(\alpha_l^n\right)^2 (t) = \sup_{[0,T]} \|\boldsymbol{w}^n(t)\|_{L^2(\Omega)}^2 = \|\boldsymbol{w}\|_{L^\infty(0,T;L^2(\Omega))}^2 .$$

The linear combination of the equations of (8.84) yields

$$(\partial_t \boldsymbol{w}^n, \boldsymbol{w}^n) + ((\nu + \nu_{\mathrm{Sm}} \|\boldsymbol{w}^n\|_{\mathrm{F}}) \nabla \boldsymbol{w}^n, \nabla \boldsymbol{w}^n) = (\boldsymbol{f}, \boldsymbol{w}^n), \tag{8.86}$$

where $n(\boldsymbol{w}^n, \boldsymbol{w}^n, \boldsymbol{w}^n) = \boldsymbol{0}$ has been used. This equation has the same form as (8.75). Since $\boldsymbol{w}^n$ is defined as a solution of (8.84), it solves also (8.86) such that the techniques used for proving Lemma 8.76 can be applied. Thus, one obtains the estimates

$$\|\boldsymbol{w}^n(t)\|_{L^2(\Omega)} \le \|\boldsymbol{w}_0\|_{L^2(\Omega)} + \left\|\boldsymbol{f}\left(t'\right)\right\|_{L^1(0,t;L^2(\Omega))}, \quad 0 \le t \le T,$$

uniformly in n, which provides the a priori boundedness which shows that there is no blow-up and thus the existence of a unique solution of (8.84) in $(0, T]$ is proved.

Analogously to the proof of Lemma 8.77, one gets

$$\|\boldsymbol{w}^n(T)\|_{L^2(\Omega)}^2 + 2\int_0^T ((\nu + \nu_{\mathrm{Sm}} \|\nabla \boldsymbol{w}^n\|_{\mathrm{F}}) \nabla \boldsymbol{w}^n, \nabla \boldsymbol{w}^n) \ dt \le C_1(T) \tag{8.87}$$

uniformly in n. System (8.84) can be brought also in form (8.79) by multiplication with $d\alpha_l^n(t) / dt$ and summation, such that the estimate

$$\|\nabla \boldsymbol{w}^n(T)\|_{L^3(\Omega)}^3 + \frac{3}{2\nu_{\mathrm{Sm}}} \|\partial_t \boldsymbol{w}^n\|_{L^2(0,T;L^2(\Omega))}^2 \le C_2(T) \tag{8.88}$$

is valid for $T > 0$ and uniformly in n. From estimates (8.87) and (8.88), one derives that $\boldsymbol{w}^n \in V$, uniformly in n. ■

Lemma 8.81 (Existence of a Converging Subsequence) *There is a function* $\boldsymbol{w} \in V$ *such that a subsequence of* $\{\boldsymbol{w}^n\}_{n=1}^{\infty}$

i) converges weakly to $\boldsymbol{w}$ *in* V,
ii) converges strongly to $\boldsymbol{w}$ *in* $L^2(0, T; L^2(\Omega))$,

iii) converges strongly to $\boldsymbol{w}$ *in* $L^q(0,T;L^q(\Omega))$ *for* $q<4$.

iv) A subsequence of $\{\partial_t \boldsymbol{w}^n\}$ *converges weakly to* $\partial_t \boldsymbol{w}$ *in* $L^2\left(0,T;L^2(\Omega)\right)$.

v) There is a subsequence of $\frac{\partial w_j^n}{\partial x_i}$, $i,j=1,\ldots,d$, *that converges weakly to* $\frac{\partial w_j}{\partial x_i}$ *in* $L^3\left(0,T;L^3(\Omega)\right)$.

Proof For brevity, it will be spoken of the convergence of $\{\boldsymbol{w}^n\}=\{\boldsymbol{w}^n\}_{n=1}^{\infty}$ instead of the convergence of a subsequence. Besides proving the convergence in the senses given in i)–v), one has to show that all kinds of convergence lead to the same limit $\boldsymbol{w}$.

i) The weak convergence of $\{\boldsymbol{w}^n\}$ to $\boldsymbol{w}\in V$, i.e.,

$$\lim_{n\to\infty}\int_0^T\int_\Omega \boldsymbol{w}^n\boldsymbol{v}\,d\boldsymbol{x}\,dt=\int_0^T\int_\Omega \boldsymbol{w}\boldsymbol{v}\,d\boldsymbol{x}\,dt\quad\forall\ \boldsymbol{v}\in V', \tag{8.89}$$

where V' is the dual space of V, follows from the uniform boundedness of $\{\boldsymbol{w}^n\}$ in the norm $\|\cdot\|_V$ of V, which is a consequence of (8.88), and that every bounded sequence in a reflexive Banach space has a weakly convergent subsequence, see Remark A.58.

ii) Since $V\subset L^2\left(0,T;L^2(\Omega)\right)\subset V'$, (8.89) holds also for all $\boldsymbol{v}\in L^2\left(0,T;L^2(\Omega)\right)$ such that $\{\boldsymbol{w}^n\}$ converges weakly to $\boldsymbol{w}$ in $L^2\left(0,T;L^2(\Omega)\right)$. From the uniform boundedness of $\boldsymbol{w}^n(t,\boldsymbol{x})$ in $W_0^{1,3}(\Omega)$ for every $t\ge 0$, estimate (8.88), and the compact imbedding $W_0^{1,3}(\Omega)$ into $L^2(\Omega)$, (A.16) and Theorem A.42 vii), it follows that there is a subsequence of $\{\boldsymbol{w}^n\}$ which converges strongly in $L^2(\Omega)$ to $\tilde{\boldsymbol{w}}\in L^2(\Omega)$. Since this statement is true for all $t\ge 0$, $\{\boldsymbol{w}^n\}$ converges strongly to $\tilde{\boldsymbol{w}}\in L^2\left(0,T;L^2(\Omega)\right)$. Consequently, this subsequence of $\{\boldsymbol{w}^n\}$ converges also weakly to $\tilde{\boldsymbol{w}}$

$$\lim_{n\to\infty}\int_0^T\int_\Omega \boldsymbol{w}^n\boldsymbol{v}\,d\boldsymbol{x}\,dt=\int_0^T\int_\Omega \tilde{\boldsymbol{w}}\boldsymbol{v}\,d\boldsymbol{x}\,dt\quad\forall\ \boldsymbol{v}\in L^2\left(0,T;L^2(\Omega)\right)\subset V'.$$

Since the weak limit is unique, it follows with (8.89) that $\boldsymbol{w}=\tilde{\boldsymbol{w}}$. Hence, (a subsequence of) $\{\boldsymbol{w}^n\}$ converges to $\boldsymbol{w}$ strongly in $L^2\left(0,T;L^2(\Omega)\right)$.

iii) It will be first established that $\{\boldsymbol{w}^n\}$ is uniformly bounded (with respect to n) in $L^4\left(0,T;L^4(\Omega)\right)$. From the Cauchy–Schwarz inequality (A.10), the Sobolev imbedding $H^1(\Omega)\to L^6(\Omega)$, see (A.22), and Poincaré's inequality (A.12), it follows that

$$\begin{aligned}\|\boldsymbol{w}^n\|_{L^4(\Omega)}^4&=\int_\Omega\|\boldsymbol{w}^n\|_2^4\,d\boldsymbol{x}\le\left(\int_\Omega\|\boldsymbol{w}^n\|_2^6\,d\boldsymbol{x}\right)^{1/2}\left(\int_\Omega\|\boldsymbol{w}^n\|_2^2\,d\boldsymbol{x}\right)^{1/2}\\&\le\|\boldsymbol{w}^n\|_{L^6(\Omega)}^3\|\boldsymbol{w}^n\|_{L^2(\Omega)}\le C\|\boldsymbol{w}^n\|_{H^1(\Omega)}^3\|\boldsymbol{w}^n\|_{L^2(\Omega)}\\&\le C\|\nabla\boldsymbol{w}^n\|_{L^2(\Omega)}^3\|\boldsymbol{w}^n\|_{L^2(\Omega)}.\end{aligned}$$

One obtains, using (8.87) and (8.88)

$$\|w^n\|^4_{L^4(0,T;L^4(\Omega))} \le C \int_0^T \|\nabla w^n\|^3_{L^2(\Omega)} \|w^n\|_{L^2(\Omega)}\ dt \le \int_0^T C_2(t)\, C_1(t)^{1/2}\ dt,$$

uniformly in n. Now, one can prove the strong convergence $w^n \to w$ in $L^q(0,T;L^q(\Omega))$, $q = 4-\varepsilon, \varepsilon > 0$. The generalized Hölder inequality (6.33) with respect to the time-space norm gives

$$\begin{aligned}
&\|w - w^n\|^q_{L^q(0,T;L^q(\Omega))} \\
&= \int_0^T \int_\Omega \|w - w^n\|_2^{2-\varepsilon} \|w - w^n\|_2 \|w - w^n\|_2\ dx\, dt \\
&\le \left\| (w - w^n)^{2-\varepsilon} \right\|_{L^2(0,T;L^2(\Omega))} \|w - w^n\|_{L^4(0,T;L^4(\Omega))} \|w - w^n\|_{L^4(0,T;L^4(\Omega))}.
\end{aligned}$$

The last two factors are bounded by the triangle inequality and the uniform boundedness of $\{w^n\}$ in $L^4(0,T;L^4(\Omega))$. The first term can be written in the form

$$\left\| (w - w^n)^{2-\varepsilon} \right\|^2_{L^2(0,T;L^2(\Omega))} = \int_0^T \int_\Omega \|w - w^n\|_2^{2-2\varepsilon} \|w - w^n\|_2 \|w - w^n\|_2\ dx\, dt.$$

Applying the same steps as for the previous estimate gives

$$\|w - w^n\|^q_{L^q(0,T;L^q(\Omega))} \le \left\| (w - w^n)^{2-2\varepsilon} \right\|_{L^2(0,T;L^2(\Omega))} \|w - w^n\|^2_{L^4(0,T;L^4(\Omega))}.$$

Continuing this approach, i.e., applying several times Hölder's inequality, yields eventually the factor

$$\left\| (w - w^n)^{2-\varepsilon_0} \right\|_{L^2(0,T;L^2(\Omega))} \quad \text{with } 2 - \varepsilon_0 \le 1.$$

If $2 - \varepsilon_0 = 1$, the result of ii) can be applied directly to prove iii). In the case $2 - \varepsilon_0 < 1$, first (A.8) has to be used to obtain

$$\left\| (w - w^n)^{2-\varepsilon_0} \right\|_{L^2(0,T;L^2(\Omega))} \le C \|w - w^n\|_{L^2(0,T;L^2(\Omega))},$$

before ii) can be used to prove iii).

iv) For $\boldsymbol{\phi} \in C_0^\infty(0,T;L^2(\Omega))$, it follows, using ii), that

$$\begin{aligned}
\lim_{n\to\infty} \int_\Omega \int_0^T \partial_t w^n \boldsymbol{\phi}\ dt\, dx &= -\lim_{n\to\infty} \int_\Omega \int_0^T w^n \partial_t \boldsymbol{\phi}\ dt\, dx \\
&= -\int_\Omega \int_0^T w \partial_t \boldsymbol{\phi}\ dt\, dx = \int_\Omega \int_0^T \partial_t w \boldsymbol{\phi}\ dt\, dx.
\end{aligned}$$

Now, statement iv) is a consequence of the density of $C_0^\infty\left(0,T;L^2\left(\Omega\right)\right)$ in $L^2\left(0,T;L^2\left(\Omega\right)\right)$, see Theorem A.38.

v) One has to prove

$$\lim_{n\to\infty}\int_0^T\int_\Omega \frac{\partial w_j^n}{\partial x_i}\phi\ d\boldsymbol{x}\ dt = \int_0^T\int_\Omega \frac{\partial w_j}{\partial x_i}\phi\ d\boldsymbol{x}\ dt \quad \forall\ \phi\in L^{3/2}\left(0,T;L^{3/2}\left(\Omega\right)\right).$$

It suffices to prove this relation for a dense set in $L^{3/2}\left(0,T;L^{3/2}\left(\Omega\right)\right)$, e.g., for functions from the space $C_0\left(0,T;C_0^1\left(\Omega\right)\right)$ that is dense in $L^{3/2}\left(0,T;L^{3/2}\left(\Omega\right)\right)$, which follows from Theorem A.38. Let $\phi\in C_0\left(0,T;C_0^1\left(\Omega\right)\right)$ be arbitrary, then applying ii) and twice integration by parts yields

$$\begin{aligned}\lim_{n\to\infty}\int_0^T\int_\Omega \frac{\partial w_j^n}{\partial x_i}\phi\ d\boldsymbol{x}\ dt &= -\lim_{n\to\infty}\int_0^T\int_\Omega w_j^n\frac{\partial\phi}{\partial x_i}\ d\boldsymbol{x}\ dt\\ &= -\int_0^T\int_\Omega w_j\frac{\partial\phi}{\partial x_i}\ d\boldsymbol{x}\ dt = \int_0^T\int_\Omega \frac{\partial w_j}{\partial x_i}\phi\ d\boldsymbol{x}\ dt.\end{aligned}$$

All boundary integrals vanish since ϕ vanishes on the boundary. Now, the proof concludes by applying ii) since it is $C_0\left(0,T;C_0^1\left(\Omega\right)\right)\subset L^2\left(0,T;L^2\left(\Omega\right)\right)$. ■

Remark 8.82 (Formulation of the Finite-Dimensional Problem with Arbitrary Test Function) In the following, the notation that "$\{\boldsymbol{w}^n\}$ converges" will be used instead of that a subsequence converges.

Let

$$P^n = \left\{\boldsymbol{v}:\boldsymbol{v}=\sum_{l=1}^n\tilde{\alpha}_l^n\left(t\right)\boldsymbol{v}_l^n\left(\boldsymbol{x}\right)\right\},$$

where $\tilde{\alpha}_l^n\left(t\right)$ are absolutely continuous functions of $t\in[0,T]$ with $\tilde{\alpha}_l^n\left(t\right)\in H^1\left(0,T\right)$, see Definition A.48. Choosing a fixed function $\boldsymbol{\phi}\in P^n$, it follows from (8.84), by taking a linear combination, that $\boldsymbol{\phi}$ satisfies

$$\begin{aligned}\int_0^T\left(\partial_t\boldsymbol{w}^n,\boldsymbol{\phi}\right)+\left(\left(\nu+\nu_{\mathrm{Sm}}\left\|\nabla\boldsymbol{w}^n\right\|_{\mathrm{F}}\right)\nabla\boldsymbol{w}^n,\nabla\boldsymbol{\phi}\right)+n\left(\boldsymbol{w}^n,\boldsymbol{w}^n,\boldsymbol{\phi}\right)\ dt\\ =\int_0^T\left(\boldsymbol{f},\boldsymbol{\phi}\right)\ dt.\end{aligned} \tag{8.90}$$

□

Lemma 8.83 (Convergence of a Bilinear Form with the Temporal Derivative and the Nonlinear Convective Term) *For $\boldsymbol{\phi}\in P^n$ it holds*

$$\lim_{n\to\infty}\int_0^T\left(\partial_t\boldsymbol{w}^n,\boldsymbol{\phi}\right)\ dt=\int_0^T\left(\partial_t\boldsymbol{w},\boldsymbol{\phi}\right)\ dt, \tag{8.91}$$

$$\lim_{n\to\infty}\int_0^T n\left(\boldsymbol{w}^n,\boldsymbol{w}^n,\boldsymbol{\phi}\right)\ dt=\int_0^T n\left(\boldsymbol{w},\boldsymbol{w},\boldsymbol{\phi}\right)\ dt. \tag{8.92}$$

Proof Equation (8.91) : This statement follows immediately from Lemma 8.81, iv).
Equation (8.92) : It is

$$\int_0^T n\left(\boldsymbol{w}^n, \boldsymbol{w}^n, \boldsymbol{\phi}\right)\, dt = \sum_{i,j=1}^{d} \int_0^T \int_\Omega w_i^n \frac{\partial w_j^n}{\partial x_i} \phi_j \, d\boldsymbol{x}\, dt.$$

Considering an arbitrary term of this sum yields

$$\int_0^T \int_\Omega w_i^n \frac{\partial w_j^n}{\partial x_i} \phi_j \, d\boldsymbol{x}\, dt = \int_0^T \int_\Omega w_i \frac{\partial w_j^n}{\partial x_i} \phi_j \, d\boldsymbol{x}\, dt + \int_0^T \int_\Omega \left(w_i^n - w_i\right) \frac{\partial w_j^n}{\partial x_i} \phi_j \, d\boldsymbol{x}\, dt.$$

Since $\frac{\partial w_j^n}{\partial x_i} \in L^3\left(0, T; L^3\left(\Omega\right)\right)$ converges weakly to $\frac{\partial w_j}{\partial x_i}$ in this space and one has from the Cauchy–Schwarz inequality (A.10) that with $w_i, \phi_j \in L^3\left(0, T; L^3\left(\Omega\right)\right)$ also $w_i\phi_j \in L^{3/2}\left(0, T; L^{3/2}\left(\Omega\right)\right)$, one obtains for the first term with Lemma 8.81 v)

$$\lim_{n\to\infty} \int_0^T \int_\Omega w_i \frac{\partial w_j^n}{\partial x_i} \phi_j \, d\boldsymbol{x}\, dt = \int_0^T \int_\Omega w_i \frac{\partial w_j}{\partial x_i} \phi_j \, d\boldsymbol{x}\, dt.$$

Applying Hölder's inequality (A.9) to the second term gives

$$\begin{aligned} &\left|\int_0^T \int_\Omega \left(w_i^n - w_i\right) \frac{\partial w_j^n}{\partial x_i} \phi_j \, d\boldsymbol{x}\, dt\right| \\ &\le \left\|\left(w_i^n - w_i\right)\phi_j\right\|_{L^{3/2}\left(0,T;L^{3/2}(\Omega)\right)} \left\|\frac{\partial w_j^n}{\partial x_i}\right\|_{L^3\left(0,T;L^3(\Omega)\right)}. \end{aligned}$$

By the Cauchy–Schwarz inequality, it follows that

$$\left\|\left(w_i^n - w_i\right)\phi_j\right\|_{L^{3/2}\left(0,T;L^{3/2}(\Omega)\right)} \le \left\|w_i^n - w_i\right\|_{L^3\left(0,T;L^3(\Omega)\right)}^{3/2} \left\|\phi_j\right\|_{L^3\left(0,T;L^3(\Omega)\right)}^{3/2}.$$

By Lemma 8.81 iii), one has that w_i^n converges strongly to w_i in $L^3\left(0, T; L^3\left(\Omega\right)\right)$. The term $\left\|\frac{\partial w_j^n}{\partial x_i}\right\|_{L^3\left(0,T;L^3(\Omega)\right)}$ is uniformly bounded and $\left\|\phi_j\right\|_{L^3\left(0,T;L^3(\Omega)\right)}^{3/2}$ is just a constant. Altogether, one gets

$$\lim_{n\to\infty} \left|\int_0^T \int_\Omega \left(w_i^n - w_i\right) \frac{\partial w_j^n}{\partial x_i} \phi_j \, d\boldsymbol{x}\, dt\right| = 0$$

and

$$\lim_{n\to\infty} \int_0^T \int_\Omega w_i^n \frac{\partial w_j^n}{\partial x_i} \phi_j \, d\boldsymbol{x}\, dt = \int_0^T \int_\Omega w_i \frac{\partial w_j}{\partial x_i} \phi_j \, d\boldsymbol{x}\, dt.$$

∎

Lemma 8.84 (The Limiting Equation) *The limiting equation of* (8.90) *is*

$$\int_0^T (\partial_t \boldsymbol{w}, \boldsymbol{\phi}) + (\boldsymbol{B}, \nabla \boldsymbol{\phi}) + n(\boldsymbol{w}, \boldsymbol{w}, \boldsymbol{\phi})\ dt = \int_0^T (\boldsymbol{f}, \boldsymbol{\phi})\ dt \tag{8.93}$$

with $\boldsymbol{B} \in L^{3/2}\left(0, T; L^{3/2}(\Omega)\right)$.

Proof The limits of the first and the third term were established in Lemma 8.83.

From the choice of $\boldsymbol{\phi}$, it follows that $\nabla \boldsymbol{\phi} \in L^3\left(0, T; L^3(\Omega)\right)$. If one can show now that the sequence $\{(\nu + \nu_{\mathrm{Sm}} \|\nabla \boldsymbol{w}^n\|_{\mathrm{F}}) \nabla \boldsymbol{w}^n\}$ is uniformly bounded in $L^{3/2}\left(0, T; L^{3/2}(\Omega)\right)$, then there is a subsequence that converges weakly to an operator $\boldsymbol{B} \in L^{3/2}\left(0, T; L^{3/2}(\Omega)\right)$, i.e.,

$$\lim_{n \to \infty} \int_0^T \left((\nu + \nu_{\mathrm{Sm}} \|\nabla \boldsymbol{w}^n\|_{\mathrm{F}}) \nabla \boldsymbol{w}^n, \nabla \boldsymbol{\phi}\right)\ dt = \int_0^T (\boldsymbol{B}, \nabla \boldsymbol{\phi})\ dt.$$

Consider the nonlinear viscous term component by component. It is

$$\begin{aligned} &\left\| (\nu + \nu_{\mathrm{Sm}} \|\nabla \boldsymbol{w}^n\|_{\mathrm{F}}) \frac{\partial w_i^n}{\partial x_j} \right\|^{3/2}_{L^{3/2}\left(0,T;L^{3/2}(\Omega)\right)} \\ &= \nu^{3/2} \int_0^T \int_\Omega \left| \frac{\partial w_i^n}{\partial x_j} \right|^{3/2} d\boldsymbol{x}\ dt + \nu_{\mathrm{Sm}}^{3/2} \int_0^T \int_\Omega \|\nabla \boldsymbol{w}^n\|_{\mathrm{F}}^{3/2} \left| \frac{\partial w_i^n}{\partial x_j} \right|^{3/2} d\boldsymbol{x}\ dt, \end{aligned} \tag{8.94}$$

$i, j = 1, \ldots, d$. Using Young's inequality (A.5) with $p = 4/3, q = 4$ gives for the first term

$$\int_0^T \int_\Omega \left| \frac{\partial w_i^n}{\partial x_j} \right|^{3/2} d\boldsymbol{x}\ dt \le \frac{3}{4} \int_0^T \int_\Omega \left| \frac{\partial w_i^n}{\partial x_j} \right|^2 d\boldsymbol{x}\ dt + \frac{1}{4} \int_0^T \int_\Omega d\boldsymbol{x}\ dt.$$

The right-hand side of this estimate is bounded by (8.87) and since Ω is bounded. For estimating the second term of (8.94), again Young's inequality with $p = 4, q = 4/3$ is used

$$\begin{aligned} &\int_0^T \int_\Omega \|\nabla \boldsymbol{w}^n\|_{\mathrm{F}}^{3/2} \left| \frac{\partial w_i^n}{\partial x_j} \right|^{3/2} d\boldsymbol{x}\ dt \\ &= \int_0^T \int_\Omega \|\nabla \boldsymbol{w}^n\|_{\mathrm{F}}^{3/4} \|\nabla \boldsymbol{w}^n\|_{\mathrm{F}}^{3/4} \left| \frac{\partial w_i^n}{\partial x_j} \right|^{3/2} d\boldsymbol{x}\ dt \\ &\le \frac{1}{4} \int_0^T \int_\Omega \|\nabla \boldsymbol{w}^n\|_{\mathrm{F}}^3\ d\boldsymbol{x}\ dt + \frac{3}{4} \int_0^T \int_\Omega \|\nabla \boldsymbol{w}^n\|_{\mathrm{F}} \left| \frac{\partial w_i^n}{\partial x_j} \right|^2 d\boldsymbol{x}\ dt \\ &= \frac{1}{4} \int_0^T \left(\|\nabla \boldsymbol{w}^n\|_{\mathrm{F}} \nabla \boldsymbol{w}^n, \nabla \boldsymbol{w}^n\right)\ dt + \frac{3}{4} \int_0^T \int_\Omega \|\nabla \boldsymbol{w}^n\|_{\mathrm{F}} \left| \frac{\partial w_i^n}{\partial x_j} \right|^2 d\boldsymbol{x}\ dt. \end{aligned}$$

Both terms in the last line of this inequality possess terms of $\nabla \boldsymbol{w}^n$ to the third power and thus they are bounded uniformly by (8.87).

Hence, (8.94) is bounded uniformly and a subsequence of $\{(\nu + \nu_{\mathrm{Sm}} \|\nabla \boldsymbol{w}^n\|_{\mathrm{F}}) \nabla \boldsymbol{w}^n\}$ converges weakly to some $\boldsymbol{B}(t, \boldsymbol{x})$. ■

Remark 8.85 (On the Limiting Equation) The limiting Eq. (8.93) holds for $\boldsymbol{\phi} \in P^n$ for arbitrary n and thus for $\boldsymbol{\phi} \in \cup_{n=1}^{\infty} P^n$. One still has to show that it holds for all test functions from V. □

Lemma 8.86 (The Limiting Equation in V**)** *The limiting Eq.* (8.93) *is valid for* $\boldsymbol{\phi} \in V$.

Proof For the proof of this lemma it is referred to Ladyzhenskaya (1969, p. 159). In this proof, it is shown that any $\boldsymbol{\phi} \in V$ can be approximated by a linear combination of functions from P^n such that the linear combination and its gradient converge to $\boldsymbol{\phi}$ and $\nabla \boldsymbol{\phi}$, respectively, in $L^2\left((0, T); L^2(\Omega)\right)$. ■

Remark 8.87 (The Nonlinear Viscous Operator) The nonlinear viscous operator is defined by $\boldsymbol{A} : L^3(\Omega) \to L^{3/2}(\Omega)$ with

$$\boldsymbol{A}(\nabla \boldsymbol{w}) = (\nu + \nu_{\mathrm{Sm}} \|\nabla \boldsymbol{w}\|_{\mathrm{F}}) \nabla \boldsymbol{w}. \tag{8.95}$$

For proving the existence of a weak solution, it will be shown that the nonlinear viscous term defines a so-called monotone operator. □

Lemma 8.88 (Strong Monotonicity of the Nonlinear Viscous Operator) *For arbitrary functions* $\boldsymbol{w}', \boldsymbol{w}'' \in W^{1,3}(\Omega)$*, it holds the estimate*

$$\int_{\Omega} \left(\boldsymbol{A}(\nabla \boldsymbol{w}') - \boldsymbol{A}(\nabla \boldsymbol{w}'')\right) : \left(\nabla \boldsymbol{w}' - \nabla \boldsymbol{w}''\right) \, d\boldsymbol{x} \geq \nu \left\|\nabla \boldsymbol{w}' - \nabla \boldsymbol{w}''\right\|_{L^2(\Omega)}^2 \tag{8.96}$$

with the operator $\boldsymbol{A}$ *defined in* (8.95)*. In addition, it is*

$$\begin{aligned}\int_{\Omega} \nu_{\mathrm{Sm}} \left(\left\|\nabla \boldsymbol{w}'\right\|_{\mathrm{F}} \nabla \boldsymbol{w}' - \left\|\nabla \boldsymbol{w}''\right\|_{\mathrm{F}} \nabla \boldsymbol{w}''\right) : \left(\nabla \boldsymbol{w}' - \nabla \boldsymbol{w}''\right) \, d\boldsymbol{x} \\ \geq \frac{\nu_{\mathrm{Sm}}}{4} \left\|\nabla \boldsymbol{w}' - \nabla \boldsymbol{w}''\right\|_{L^3(\Omega)}^3,\end{aligned} \tag{8.97}$$

i.e., the Smagorinsky term defines a strongly monotone operator from $L^3(\Omega)$ *into* $L^{3/2}(\Omega)$.

Proof Let $\boldsymbol{w}', \boldsymbol{w}'' \in C^1\left(\overline{\Omega}\right)$. With $\boldsymbol{w}^\tau = \tau \boldsymbol{w}' + (1-\tau)\,\boldsymbol{w}''$, one obtains, applying the fundamental theorem of calculus,

$$\begin{aligned}
&\left(\boldsymbol{A}\left(\nabla \boldsymbol{w}'\right) - \boldsymbol{A}\left(\nabla \boldsymbol{w}''\right)\right) : \left(\nabla \boldsymbol{w}' - \nabla \boldsymbol{w}''\right) \\
&= \sum_{i,j=1}^{d} \left(\boldsymbol{A}_{ij}\left(\nabla \boldsymbol{w}'\right) - \boldsymbol{A}_{ij}\left(\nabla \boldsymbol{w}''\right)\right)\left(\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right) \\
&= \sum_{i,j=1}^{d} \left(\int_0^1 \frac{d}{d\tau}\boldsymbol{A}_{ij}\left(\nabla \boldsymbol{w}^\tau\right) d\tau\right)\left(\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right).
\end{aligned} \tag{8.98}$$

Applying the product rule, one gets

$$\frac{d}{d\tau}\boldsymbol{A}_{ij}\left(\nabla \boldsymbol{w}^\tau\right) = \left(\nu_{\mathrm{Sm}} \frac{\partial}{\partial \tau}\left\|\nabla \boldsymbol{w}^\tau\right\|_{\mathrm{F}}\right)\frac{\partial w_i^\tau}{\partial x_j} + \left(\nu + \nu_{\mathrm{Sm}}\left\|\nabla \boldsymbol{w}^\tau\right\|_{\mathrm{F}}\right)\left(\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right).$$

With the chain rule, one obtains

$$\begin{aligned}
\frac{\partial}{\partial \tau}\left\|\nabla \boldsymbol{w}^\tau\right\|_{\mathrm{F}} &= \frac{\partial}{\partial \tau}\left(\sum_{k,l=1}^{d}\left(\tau \frac{\partial w_k'}{\partial x_l} + (1-\tau)\frac{\partial w_k''}{\partial x_l}\right)^2\right)^{1/2} \\
&= \frac{1}{2}\frac{\sum_{k,l=1}^{d} 2\left(\tau \frac{\partial w_k'}{\partial x_l} + (1-\tau)\frac{\partial w_k''}{\partial x_l}\right)\left(\frac{\partial w_k'}{\partial x_l} - \frac{\partial w_k''}{\partial x_l}\right)}{\left(\sum_{k,l=1}^{d}\left(\tau \frac{\partial w_k'}{\partial x_l} + (1-\tau)\frac{\partial w_k''}{\partial x_l}\right)^2\right)^{1/2}} \\
&= \frac{1}{\left\|\nabla \boldsymbol{w}^\tau\right\|_{\mathrm{F}}}\sum_{k,l=1}^{d}\frac{\partial w_k^\tau}{\partial x_l}\left(\frac{\partial w_k'}{\partial x_l} - \frac{\partial w_k''}{\partial x_l}\right).
\end{aligned}$$

Inserting these expressions in (8.98) yields

$$\begin{aligned}
&\left(\boldsymbol{A}\left(\nabla \boldsymbol{w}'\right) - \boldsymbol{A}\left(\nabla \boldsymbol{w}''\right)\right) : \left(\nabla \boldsymbol{w}' - \nabla \boldsymbol{w}''\right) \\
&= \int_0^1 \sum_{i,j=1}^{d}\left(\nu + \nu_{\mathrm{Sm}}\left\|\nabla \boldsymbol{w}^\tau\right\|_{\mathrm{F}}\right)\left(\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right)\left(\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right) d\tau \\
&\quad + \int_0^1 \nu_{\mathrm{Sm}}\left\|\nabla \boldsymbol{w}^\tau\right\|_{\mathrm{F}}^{-1}\sum_{i,j,k,l=1}^{d}\frac{\partial w_k^\tau}{\partial x_l}\frac{\partial w_i^\tau}{\partial x_j}\left(\frac{\partial w_k'}{\partial x_l} - \frac{\partial w_k''}{\partial x_l}\right)\left(\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right) d\tau.
\end{aligned} \tag{8.99}$$

The second term is non-negative since

$$\sum_{i,j,k,l=1}^{d}\frac{\partial w_k^\tau}{\partial x_l}\frac{\partial w_i^\tau}{\partial x_j}\left(\frac{\partial w_k'}{\partial x_l} - \frac{\partial w_k''}{\partial x_l}\right)\left(\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right) = \left(\sum_{i,j=1}^{d}\frac{\partial w_i^\tau}{\partial x_j}\left(\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right)\right)^2. \tag{8.100}$$

Thus, this term can be estimated from below by zero.

Estimate (8.96) The first term of (8.99) is estimated by using the non-negativity of the term including the turbulent viscosity

$$\begin{aligned}
&\int_0^1 \sum_{i,j,=1}^{d} \left(\nu + \nu_{\mathrm{Sm}} \left\|\nabla \boldsymbol{w}^\tau\right\|_{\mathrm{F}}\right) \left(\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right)^2 d\tau \\
&\geq \sum_{i,j,=1}^{d} \nu \left(\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right)^2 \int_0^1 d\tau \\
&= \nu \left(\nabla \boldsymbol{w}' - \nabla \boldsymbol{w}''\right) : \left(\nabla \boldsymbol{w}' - \nabla \boldsymbol{w}''\right) = \nu \left\|\nabla \boldsymbol{w}' - \nabla \boldsymbol{w}''\right\|_{\mathrm{F}}^2 .
\end{aligned} \tag{8.101}$$

Estimate (8.97) For this estimate, the term with ν in (8.99) is bounded by zero from below. The estimate of the other term starts by bounding the Frobenius norm from below by the largest term of the sum and using the definition of $\boldsymbol{w}_k^\tau$

$$\begin{aligned}
&\int_0^1 \sum_{i,j=1}^{d} \nu_{\mathrm{Sm}} \left\|\nabla \boldsymbol{w}^\tau\right\|_{\mathrm{F}} \left(\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right)^2 \, d\tau \\
&\geq \nu_{\mathrm{Sm}} \int_0^1 \sum_{i,j=1}^{d} \max_{k,l=1,\ldots,d} \left|\frac{\partial w_k^\tau}{\partial x_l}\right| \left(\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right)^2 \, d\tau \\
&= \nu_{\mathrm{Sm}} \int_0^1 \sum_{i,j=1}^{d} \max_{k,l=1,\ldots,d} \left|\tau \frac{\partial w_k'}{\partial x_l} + (1-\tau) \frac{\partial w_k''}{\partial x_l}\right| \left(\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right)^2 \, d\tau .
\end{aligned} \tag{8.102}$$

Now, the estimate

$$\int_0^1 \left|\tau a + (1-\tau) b\right| \, d\tau \geq \frac{|a-b|}{4}, \quad a, b \in \mathbb{R},$$

will be applied. To prove this estimate, one has to distinguish the cases $ab \geq 0$ and $ab < 0$. In the first case, one finds, e.g., for $a, b \leq 0$, with a straightforward calculation of the integral and using the triangle inequality

$$\begin{aligned}
\int_0^1 \left|\tau a + (1-\tau) b\right| \, d\tau &= -\int_0^1 \tau a + (1-\tau) b \, d\tau = -\frac{a+b}{2} = \frac{|a|+|b|}{2} \geq \frac{|a|+|b|}{4} \\
&\geq \frac{|a-b|}{4}.
\end{aligned}$$

The calculation for $a, b \geq 0$ is performed analogously. In the second case, one finds that the term in the integral, which is linear in τ, has a root at $\tau_0 = -b/(a-b)$. Now, one splits the integral into the integrals on $(0, \tau_0)$ and $(\tau_0, 1)$. Consider, e.g., the case $a > 0$ and $b < 0$, one finds with a direct calculation

$$\int_0^1 |\tau a + (1-\tau)b| \ d\tau = -\int_0^{\tau_0} (\tau a + (1-\tau)b) \ d\tau + \int_0^1 (\tau a + (1-\tau)b) \ d\tau$$
$$= a\left(\frac{1}{2} - \tau_0^2\right) + b\left(-\frac{1}{2} + (1-\tau_0)^2\right) = \frac{a^2+b^2}{2(a-b)}.$$

Then, dropping a non-negative term yields

$$\int_0^1 |\tau a + (1-\tau)b| \ d\tau = \frac{1}{2}\frac{a^2+b^2}{a-b} = \frac{1}{4}\frac{(a-b)^2+(a+b)^2}{a-b} \geq \frac{a-b}{4} = \frac{|a-b|}{4}.$$

For $a < 0$ and $b > 0$, one gets the same estimate. Inserting this estimate in (8.102) and estimating the maximal value by the absolute value of the individual terms from below gives

$$\int_0^1 \sum_{i,j=1}^d \nu_{\mathrm{Sm}} \|\nabla \boldsymbol{w}^\tau\|_{\mathrm{F}} \left(\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right)^2 \ d\tau$$
$$\geq \nu_{\mathrm{Sm}} \sum_{i,j=1}^d \max_{k,l=1,\ldots,d} \frac{1}{4}\left|\frac{\partial w_k'}{\partial x_l} - \frac{\partial w_k''}{\partial x_l}\right| \left(\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right)^2$$
$$\geq \frac{\nu_{\mathrm{Sm}}}{4} \sum_{i,j=1}^d \left|\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right| \left(\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right)^2$$
$$= \frac{\nu_{\mathrm{Sm}}}{4} \sum_{i,j=1}^d \left|\frac{\partial w_i'}{\partial x_j} - \frac{\partial w_i''}{\partial x_j}\right|^3.$$

The proof of both estimates is completed with integration on Ω and using the density of $C^1\left(\overline{\Omega}\right)$ in $W^{1,3}(\Omega)$, which is a consequence of Theorem A.38. ■

Lemma 8.89 (Estimate of the Difference of the Weak Equation for Two Test Functions) *Let* $\tilde{\boldsymbol{\phi}} \in V$*, then it is*

$$-\int_0^T \int_\Omega (\partial_t \boldsymbol{w} + (\boldsymbol{w}\cdot\nabla)\boldsymbol{w} - \boldsymbol{f}) \cdot \left(\boldsymbol{w} - \tilde{\boldsymbol{\phi}}\right) + \boldsymbol{A}\left(\nabla\tilde{\boldsymbol{\phi}}\right) : \nabla\left(\boldsymbol{w} - \tilde{\boldsymbol{\phi}}\right) \ d\boldsymbol{x} \ dt \geq 0. \tag{8.103}$$

Proof To begin with, let $\tilde{\boldsymbol{\phi}} \in P^n$ for an arbitrary n. From (8.96), it follows that

$$\int_0^T \int_\Omega \left(\boldsymbol{A}(\nabla \boldsymbol{w}^n) - \boldsymbol{A}\left(\nabla\tilde{\boldsymbol{\phi}}\right)\right) : \left(\nabla \boldsymbol{w}^n - \nabla\tilde{\boldsymbol{\phi}}\right) \ d\boldsymbol{x} \ dt \geq 0. \tag{8.104}$$

Since $\boldsymbol{w}^n \in P^n$ and $\tilde{\boldsymbol{\phi}} \in P^n$, one obtains with $\boldsymbol{\phi} = \boldsymbol{w}^n - \tilde{\boldsymbol{\phi}}$ in (8.90)

$$\int_0^T \int_\Omega \boldsymbol{A}\left(\nabla \boldsymbol{w}^n\right) : \left(\nabla \boldsymbol{w}^n - \nabla \tilde{\boldsymbol{\phi}}\right)\, d\boldsymbol{x}\, dt = \int_0^T \left(\boldsymbol{f} - \partial_t \boldsymbol{w}^n - \left(\boldsymbol{w}^n \cdot \nabla\right) \boldsymbol{w}^n, \boldsymbol{w}^n - \tilde{\boldsymbol{\phi}}\right)\, dt,$$

such that with (8.90)

$$-\int_0^T \int_\Omega \Big[\left(\partial_t \boldsymbol{w}^n + \left(\boldsymbol{w}^n \cdot \nabla\right) \boldsymbol{w}^n - \boldsymbol{f}\right) \cdot \left(\boldsymbol{w}^n - \tilde{\boldsymbol{\phi}}\right) + \boldsymbol{A}\left(\nabla \tilde{\boldsymbol{\phi}}\right) : \left(\nabla \boldsymbol{w}^n - \nabla \tilde{\boldsymbol{\phi}}\right) \Big]\, d\boldsymbol{x}\, dt \geq 0. \tag{8.105}$$

It is

$$\int_0^T \left(\boldsymbol{f}, \boldsymbol{w}^n - \tilde{\boldsymbol{\phi}}\right)\, dt = \int_0^T \left(\boldsymbol{f}, \boldsymbol{w} - \tilde{\boldsymbol{\phi}}\right)\, dt + \int_0^T \left(\boldsymbol{f}, \boldsymbol{w}^n - \boldsymbol{w}\right)\, dt.$$

The second term converges to zero since $\boldsymbol{w}^n$ converges to $\boldsymbol{w}$ strongly in $L^2\left(0, T; L^2(\Omega)\right)$ and

$$\left| \int_0^T \left(\boldsymbol{f}, \boldsymbol{w}^n - \boldsymbol{w}\right)\, dt \right| \leq \|\boldsymbol{f}\|_{L^2\left(0,T;L^2(\Omega)\right)} \|\boldsymbol{w}^n - \boldsymbol{w}\|_{L^2\left(0,T;L^2(\Omega)\right)}.$$

The nonlinear term is considered component by component

$$\int_0^T \int_\Omega \left(\boldsymbol{w}^n \cdot \nabla\right) \boldsymbol{w}^n \cdot \boldsymbol{w}^n\, d\boldsymbol{x}\, dt = \sum_{k,l=1}^d \int_0^T \int_\Omega w_k^n \frac{\partial w_l^n}{\partial x_k} w_l^n\, d\boldsymbol{x}\, dt.$$

By construction, it is $\frac{\partial w_l^n}{\partial x_k} \in L^3\left(0, T; L^3(\Omega)\right)$ and it was proved in Lemma 8.81 v) that $\frac{\partial w_l^n}{\partial x_k} \rightharpoonup \frac{\partial w_l}{\partial x_k}$ in $L^3\left(0, T; L^3(\Omega)\right)$. Hence, it follows that

$$\begin{aligned}
&\lim_{n\to\infty} \int_0^T \int_\Omega w_k^n \frac{\partial w_l^n}{\partial x_k} w_l^n\, d\boldsymbol{x}\, dt \\
&= \lim_{n\to\infty} \int_0^T \int_\Omega w_k \frac{\partial w_l^n}{\partial x_k} w_l\, d\boldsymbol{x}\, dt + \lim_{n\to\infty} \int_0^T \int_\Omega \frac{\partial w_l^n}{\partial x_k} \left(w_k^n w_l^n - w_k w_l\right)\, d\boldsymbol{x}\, dt \\
&= \int_0^T \int_\Omega w_k \frac{\partial w_l}{\partial x_k} w_l\, d\boldsymbol{x}\, dt + \lim_{n\to\infty} \int_0^T \int_\Omega \frac{\partial w_l^n}{\partial x_k} \left(w_k^n w_l^n - w_k w_l\right)\, d\boldsymbol{x}\, dt.
\end{aligned}$$

Since one obtains with Hölder's inequality (A.9)

$$\begin{aligned}
&\left| \int_0^T \int_\Omega \frac{\partial w_l^n}{\partial x_k} \left(w_k^n w_l^n - w_k w_l\right)\, d\boldsymbol{x}\, dt \right| \\
&\leq \left\| \frac{\partial w_l^n}{\partial x_k} \right\|_{L^3\left(0,T;L^3(\Omega)\right)} \left\| w_k^n w_l^n - w_k w_l \right\|_{L^{3/2}\left(0,T;L^{3/2}(\Omega)\right)},
\end{aligned}$$

one has to show that $w_k^n w_l^n \to w_k w_l$ strongly in $L^{3/2}\left(0,T;L^{3/2}(\Omega)\right)$. Using the triangle inequality and the Cauchy–Schwarz inequality (A.10) gives

$$\begin{aligned}
&\left\|w_k^n w_l^n - w_k w_l\right\|_{L^{3/2}\left(0,T;L^{3/2}(\Omega)\right)} \\
&\le \left\|\left(w_k^n - w_k\right) w_l^n\right\|_{L^{3/2}\left(0,T;L^{3/2}(\Omega)\right)} + \left\|w_k\left(w_l^n - w_l\right)\right\|_{L^{3/2}\left(0,T;L^{3/2}(\Omega)\right)} \\
&\le \left\|w_k^n - w_k\right\|_{L^3\left(0,T;L^3(\Omega)\right)} \left\|w_l^n\right\|_{L^3\left(0,T;L^3(\Omega)\right)} \\
&\quad + \left\|w_l^n - w_l\right\|_{L^3\left(0,T;L^3(\Omega)\right)} \left\|w_l^k\right\|_{L^3\left(0,T;L^3(\Omega)\right)}.
\end{aligned}$$

Because $w_k^n \to w_k$ strongly in $L^3\left(0,T;L^3(\Omega)\right)$, see Lemma 8.81 iii), it follows that $w_k^n w_l^n \to w_k w_l$ strongly in $L^{3/2}\left(0,T;L^{3/2}(\Omega)\right)$ and

$$\lim_{n\to\infty} \int_0^T \int_\Omega w_k^n \frac{\partial w_l^n}{\partial x_k} l w_l^n \, d\boldsymbol{x}\, dt = \int_0^T \int_\Omega w_k \frac{\partial w_l}{\partial x_k} w_l \, d\boldsymbol{x}\, dt.$$

It was shown in the proof of Lemma 8.84 that $\boldsymbol{A}\left(\nabla\tilde{\boldsymbol{\phi}}\right) \in L^{3/2}\left(0,T;L^{3/2}(\Omega)\right)$. Since $\nabla\boldsymbol{w}^n \in L^3\left(0,T;L^3(\Omega)\right)$ and $\nabla\boldsymbol{w}^n \rightharpoonup \nabla\boldsymbol{w}$ in that space, see Lemma 8.81 v), one gets

$$\lim_{n\to\infty} \int_0^T \int_\Omega \boldsymbol{A}\left(\nabla\tilde{\boldsymbol{\phi}}\right) : \nabla\boldsymbol{w}^n \, d\boldsymbol{x}\, dt = \int_0^T \int_\Omega \boldsymbol{A}\left(\nabla\tilde{\boldsymbol{\phi}}\right) : \nabla\boldsymbol{w} \, d\boldsymbol{x}\, dt.$$

In addition, it is, using the Cauchy–Schwarz inequality

$$\begin{aligned}
&\lim_{n\to\infty} \int_0^T \int_\Omega \partial_t \boldsymbol{w}^n \boldsymbol{w}^n \, d\boldsymbol{x}\, dt \\
&= \lim_{n\to\infty} \int_0^T \int_\Omega \partial_t \boldsymbol{w}^n \left(\boldsymbol{w}^n - \boldsymbol{w}\right) \, d\boldsymbol{x}\, dt + \lim_{n\to\infty} \int_0^T \int_\Omega \partial_t \boldsymbol{w}^n \boldsymbol{w} \, d\boldsymbol{x}\, dt \\
&\le \lim_{n\to\infty} \left\|\partial_t \boldsymbol{w}^n\right\|_{L^2\left(0,T;L^2(\Omega)\right)} \left\|\boldsymbol{w}^n - \boldsymbol{w}\right\|_{L^2\left(0,T;L^2(\Omega)\right)} + \lim_{n\to\infty} \int_0^T \int_\Omega \partial_t \boldsymbol{w}^n \boldsymbol{w} \, d\boldsymbol{x}\, dt \\
&= 0 + \int_0^T \int_\Omega \partial_t \boldsymbol{w}\boldsymbol{w} \, d\boldsymbol{x}\, dt
\end{aligned}$$

by Lemma 8.81 ii) and iv).

Thus, inequality (8.103) is proved for any $\tilde{\boldsymbol{\phi}} \in P^n$ with arbitrary n. But then it is also valid for arbitrary $\tilde{\boldsymbol{\phi}} \in P$ and also for arbitrary $\tilde{\boldsymbol{\phi}} \in V$ by Lemma 8.86. ■

Lemma 8.90 (Identifying *B* and the Nonlinear Viscous Operator) *For all $\boldsymbol{\phi} \in V$, it holds*

$$\int_0^T \int_\Omega \boldsymbol{B} : \nabla\boldsymbol{\phi} \, d\boldsymbol{x}\, dt = \int_0^T \int_\Omega \boldsymbol{A}\left(\nabla\boldsymbol{w}\right) : \nabla\boldsymbol{\phi} \, d\boldsymbol{x}\, dt.$$

Proof Since (8.93) is valid for all $\boldsymbol{\phi} \in V$, one can choose $\boldsymbol{\phi} = \boldsymbol{w} - \tilde{\boldsymbol{\phi}}$. Adding (8.103) to this equation yields

$$\int_0^T \int_\Omega \left(\boldsymbol{B} - \boldsymbol{A}\left(\nabla\tilde{\boldsymbol{\phi}}\right)\right) : \left(\nabla \boldsymbol{w} - \nabla\tilde{\boldsymbol{\phi}}\right)\, d\boldsymbol{x}\, dt \geq 0.$$

Setting $\tilde{\boldsymbol{\phi}} = \boldsymbol{w} - \varepsilon\hat{\boldsymbol{\phi}}$ with $\varepsilon > 0$ and $\hat{\boldsymbol{\phi}} \in V$ arbitrary gives

$$\varepsilon \int_0^T \int_\Omega \left(\boldsymbol{B} - \boldsymbol{A}\left(\nabla \boldsymbol{w} - \varepsilon\nabla\hat{\boldsymbol{\phi}}\right)\right) : \nabla\hat{\boldsymbol{\phi}}\, d\boldsymbol{x}\, dt \geq 0.$$

Division by ε, taking the limit $\varepsilon \to 0$, and using the continuity of the Frobenius norm leads to

$$\int_0^T \int_\Omega \left(\boldsymbol{B} - \boldsymbol{A}\left(\nabla \boldsymbol{w}\right)\right) : \nabla\hat{\boldsymbol{\phi}}\, d\boldsymbol{x}\, dt \geq 0. \tag{8.106}$$

If $\hat{\boldsymbol{\phi}} \in V$, then also $-\hat{\boldsymbol{\phi}} \in V$. Thus, if the integral is positive for $\hat{\boldsymbol{\phi}}$, then it is negative for $-\hat{\boldsymbol{\phi}}$ which is a contradiction to (8.106), since (8.106) holds for all functions from V. Hence, it follows that

$$\int_0^T \int_\Omega \left(\boldsymbol{B} - \boldsymbol{A}\left(\nabla \boldsymbol{w}\right)\right) : \nabla\hat{\boldsymbol{\phi}}\, d\boldsymbol{x}\, dt = 0 \quad \forall\, \hat{\boldsymbol{\phi}} \in V,$$

which is the statement of the lemma. ■

Theorem 8.91 (Existence of a Weak Solution) *Problem* (8.73) *possesses at least one solution* $\boldsymbol{w} \in V$ *for arbitrary* $\boldsymbol{f} \in L^2\left(0, T; L^2\left(\Omega\right)\right)$ *and* $\boldsymbol{w}_0 \in W^{1,3}_{0,\mathrm{div}}\left(\Omega\right)$.

Proof It was shown, starting with Lemma 8.81, that there is a $\boldsymbol{w} \in V$ that satisfies the weak formulation of (8.73). Since $\boldsymbol{w}$ is given as a limit of a sequence $\{\boldsymbol{w}^n\}_{n=1}^\infty$ with $\boldsymbol{w}^n\left(0, \boldsymbol{x}\right) = \boldsymbol{w}_0\left(\boldsymbol{x}\right)$ for all n, it follows that $\boldsymbol{w}\left(0, \boldsymbol{x}\right) = \boldsymbol{w}_0\left(\boldsymbol{x}\right)$, such that the initial condition is satisfied, too. ■

Remark 8.92 (Generalizations)

- In Ladyženskaja (1967), the existence of a solution has been proved for the weak formulation of the form

$$\int_0^T \left(\partial_t \boldsymbol{w} + \left(\boldsymbol{w} \cdot \nabla\right) \boldsymbol{w}, \boldsymbol{v}\right) + \left(\left(\nu + \nu_{\mathrm{Sm}} \left\|\nabla \boldsymbol{w}\right\|_{\mathrm{F}}^{\mu}\right) \nabla \boldsymbol{w}, \nabla \boldsymbol{v}\right)\, dt = \int_0^T \left(\boldsymbol{f}, \boldsymbol{v}\right)\, dt$$

  with $\mu \geq 2/5$. The restriction on μ comes from the application of the Sobolev imbedding to obtain inequality (8.80).
- The result of Ladyzhenskaya was extended in Du and Gunzburger (1991) to $\mu \geq 1/5$ by deriving new stability (a priori) estimates.

- In Świerczewska (2006), the case of a non-constant parameter in the Smagorinsky model was studied. This parameter function was allowed to depend on filtered functions of $\boldsymbol{w}$ and $\mathbb{D}(\boldsymbol{w})$, see Świerczewska (2006) for details. It was assumed that the parameter function is continuous, it is bounded from below by a positive constant, and it is bounded from above. With these conditions, the existence of a weak solution was proved.

□

Theorem 8.93 (Uniqueness of the Weak Solution) *Under the same assumptions as in Theorem 8.91, the weak solution of* (8.73) *is unique in* V.

Proof Assume that there are two weak solutions $\boldsymbol{w}', \boldsymbol{w}'' \in V$ of (8.73) and denote $\tilde{\boldsymbol{w}} = \boldsymbol{w}' - \boldsymbol{w}''$. Thus, $\tilde{\boldsymbol{w}} \in V$ and $\tilde{\boldsymbol{w}}(0, \boldsymbol{x}) = \boldsymbol{0}$. Subtracting (8.73) for $\boldsymbol{w} = \boldsymbol{w}'$, $\boldsymbol{v} = \tilde{\boldsymbol{w}}$ and $\boldsymbol{w} = \boldsymbol{w}''$, $\boldsymbol{v} = \tilde{\boldsymbol{w}}$ gives

$$\begin{aligned} 0 = \int_0^T & (\partial_t \tilde{\boldsymbol{w}}, \tilde{\boldsymbol{w}}) + \left(\boldsymbol{A}\left(\nabla \boldsymbol{w}'\right) - \boldsymbol{A}\left(\nabla \boldsymbol{w}''\right), \nabla \boldsymbol{w}' - \nabla \boldsymbol{w}''\right) \\ & + n\left(\boldsymbol{w}', \boldsymbol{w}', \tilde{\boldsymbol{w}}\right) - n\left(\boldsymbol{w}'', \boldsymbol{w}'', \tilde{\boldsymbol{w}}\right) \, dt. \end{aligned}$$

Adding and subtracting $n(\boldsymbol{w}'', \boldsymbol{w}', \tilde{\boldsymbol{w}})$, using $n(\boldsymbol{w}'', \tilde{\boldsymbol{w}}, \tilde{\boldsymbol{w}}) = 0$, see Lemmas 6.10, and (7.13), this equation can be rewritten in the following form

$$0 = \int_0^T \frac{d}{dt} \|\tilde{\boldsymbol{w}}\|_{L^2(\Omega)}^2 + 2\left(\boldsymbol{A}\left(\nabla \boldsymbol{w}'\right) - \boldsymbol{A}\left(\nabla \boldsymbol{w}''\right), \nabla \boldsymbol{w}' - \nabla \boldsymbol{w}''\right) + 2n\left(\tilde{\boldsymbol{w}}, \boldsymbol{w}', \tilde{\boldsymbol{w}}\right) \, dt.$$

Using the strong monotonicity property (8.97) of $\boldsymbol{A}(\cdot)$, Hölder's inequality (A.9), the Sobolev imbedding $H^1(\Omega) \to L^6(\Omega)$, see (A.22), and Poincaré's inequality (A.12) leads to

$$\begin{aligned} & \int_0^T \frac{d}{dt} \|\tilde{\boldsymbol{w}}\|_{L^2(\Omega)}^2 + 2\nu \|\nabla \tilde{\boldsymbol{w}}\|_{L^2(\Omega)}^2 + \frac{\nu_T}{2} \|\nabla \tilde{\boldsymbol{w}}\|_{L^3(\Omega)}^3 \, dt \\ & \leq -2 \int_0^T n\left(\tilde{\boldsymbol{w}}, \boldsymbol{w}', \tilde{\boldsymbol{w}}\right) \, dt \\ & \leq 2 \int_0^T \|\tilde{\boldsymbol{w}}\|_{L^2(\Omega)} \left\|\nabla \boldsymbol{w}'\right\|_{L^3(\Omega)} \|\tilde{\boldsymbol{w}}\|_{L^6(\Omega)} \, dt \\ & \leq C \int_0^T \|\tilde{\boldsymbol{w}}\|_{L^2(\Omega)} \left\|\nabla \boldsymbol{w}'\right\|_{L^3(\Omega)} \|\nabla \tilde{\boldsymbol{w}}\|_{L^2(\Omega)} \, dt. \end{aligned} \tag{8.107}$$

The right-hand side can be estimated further by Young's inequality (A.5)

$$\begin{aligned} & \int_0^T \frac{d}{dt} \|\tilde{\boldsymbol{w}}\|_{L^2(\Omega)}^2 + 2\nu \|\nabla \tilde{\boldsymbol{w}}\|_{L^2(\Omega)}^2 + \frac{\nu_T}{2} \|\nabla \tilde{\boldsymbol{w}}\|_{L^3(\Omega)}^3 \, dt \\ & \leq \int_0^T 2\nu \|\nabla \tilde{\boldsymbol{w}}\|_{L^2(\Omega)}^2 + \frac{C^2}{8\nu} \|\tilde{\boldsymbol{w}}\|_{L^2(\Omega)}^2 \left\|\nabla \boldsymbol{w}'\right\|_{L^3(\Omega)}^2 \, dt. \end{aligned} \tag{8.108}$$

Neglecting terms on the left-hand side, integrating on $(0, T)$, and using $\tilde{\boldsymbol{w}}(0) = \boldsymbol{0}$ yields

$$\|\tilde{\boldsymbol{w}}(T)\|_{L^2(\Omega)}^2 \le C \int_0^T \|\tilde{\boldsymbol{w}}\|_{L^2(\Omega)}^2 \left\|\nabla \boldsymbol{w}'\right\|_{L^3(\Omega)}^2 \, dt. \tag{8.109}$$

Gronwall's lemma (A.34) now gives $\|\tilde{\boldsymbol{w}}(T)\|_{L^2(\Omega)}^2 \le 0$ for all T which proves the uniqueness of the solution in $L^\infty\left(0, T; L^2(\Omega)\right)$. Since it is $H^1\left(0, T; L^2(\Omega)\right) \subset L^\infty\left(0, T; L^2(\Omega)\right)$ by the Sobolev imbedding (A.18) in one dimension, $\tilde{\boldsymbol{w}}$ is in the equivalence class of zero in $H^1\left(0, T; L^2(\Omega)\right)$ and hence the solution is unique in $H^1\left(0, T; L^2(\Omega)\right)$.

Applying now the result $\|\tilde{\boldsymbol{w}}(t)\|_{L^2(\Omega)}^2 = 0$ for almost all t, which implies in particular $\frac{d}{dt}\|\tilde{\boldsymbol{w}}\|_{L^2(\Omega)}^2 = 0$ for almost all t, in (8.108) gives

$$\int_0^T \|\nabla \tilde{\boldsymbol{w}}\|_{L^3(\Omega)}^3 \, dt \le 0$$

for almost all T. Hence $\|\nabla \tilde{\boldsymbol{w}}\|_{L^3\left(0,T;L^3(\Omega)\right)} = 0$, i.e., $\tilde{\boldsymbol{w}}$ is in the equivalence class of zero in $L^3\left(0, T; W_{0,\mathrm{div}}^{1,3}(\Omega)\right)$, which proves the uniqueness of the solution in $L^3\left(0, T; W_{0,\mathrm{div}}^{1,3}(\Omega)\right)$. ■

Theorem 8.94 (Stability of the Weak Solution) *Let the assumptions of Theorem 8.91 be satisfied and let $\boldsymbol{w}', \boldsymbol{w}'' \in V$ be solutions of (8.73) with different initial data and different right-hand sides $\boldsymbol{f}', \boldsymbol{f}''$. Then it is*

$$\begin{aligned} &\left\|\boldsymbol{w}' - \boldsymbol{w}''\right\|_{L^\infty\left(0,T;L^2(\Omega)\right)}^2 \\ &\le \left(\left\|\boldsymbol{w}'(0) - \boldsymbol{w}''(0)\right\|_{L^2(\Omega)}^2 + \frac{1}{2C_1(T)}\left\|\boldsymbol{f}' - \boldsymbol{f}''\right\|_{L^2\left(0,T;L^2(\Omega)\right)}^2\right) \\ &\quad \times \exp\left(C_2\left\|\nabla \boldsymbol{w}'\right\|_{L^2\left(0,T;L^3(\Omega)\right)}^2 + \frac{C_1(T)}{2} T\right), \end{aligned}$$

with $C_1(T), C_2 > 0$ and $C_1(T)$ can be chosen arbitrarily.

If $\boldsymbol{f}' = \boldsymbol{f}''$, then it holds

$$\begin{aligned} &\left\|\boldsymbol{w}' - \boldsymbol{w}''\right\|_{L^\infty\left(0,T;L^2(\Omega)\right)}^2 \\ &\le \left\|\boldsymbol{w}'(0) - \boldsymbol{w}''(0)\right\|_{L^2(\Omega)}^2 \exp\left(C_2\left\|\nabla \boldsymbol{w}'\right\|_{L^2\left(0,T;L^3(\Omega)\right)}^2\right). \end{aligned}$$

Proof The proof starts in the same way as the proof of Theorem 8.93. One obtains instead of (8.109)

$$\|\tilde{\boldsymbol{w}}(T)\|_{L^2(\Omega)}^2 \le \|\tilde{\boldsymbol{w}}(0)\|_{L^2(\Omega)}^2 + \int_0^T \frac{1}{2C_1(T)} \|\boldsymbol{f}' - \boldsymbol{f}''\|_{L^2(\Omega)}^2 \, dt + \int_0^T \left(\frac{C^2}{4\nu} \|\nabla \boldsymbol{w}'\|_{L^3(\Omega)}^2 + \frac{C_1(T)}{2} \right) \|\tilde{\boldsymbol{w}}\|_{L^2(\Omega)}^2 \, dt.$$

Setting $C^2/(4\nu) = C_2$ and applying Gronwall's lemma (A.33) proves the statement of the theorem. ■

Remark 8.95 (Other Analytical Investigations of the Smagorinsky Model)

- Parés (1992) studied the existence and uniqueness of a weak solution of a Smagorinsky model that differs from (8.70) in some aspects. First, the deformation tensor formulation of the viscous term and the deformation tensor formulation of the Smagorinsky model, i.e., (8.67), were considered. Second, homogeneous Dirichlet boundary conditions are prescribed only at a part of the boundary Γ_{nosl} with $|\Gamma_{\text{nosl}}| > 0$. On the rest of the boundary, slip with friction and penetration with resistance boundary conditions are given, e.g., the linear conditions described in Remark 2.26. These boundary conditions lead to an additional term in the weak formulation of the momentum equation of the Smagorinsky model. The existence proof uses the Galerkin method in the same way as described in the present section. In addition, estimates for the additional term coming from the boundary conditions have to be proved.
- The Smagorinsky model can be used to stabilize the dominating convection in the stationary Navier–Stokes equations (6.1). In Parés (1992), the stationary Smagorinsky model is considered with the same features as the time-dependent model. The existence of weak solutions and the uniqueness in the case of small data could be proved. In Du and Gunzburger (1991), existence of a unique weak solution for small data was shown.

□

8.3.3 Finite Element Error Analysis for the Time-Continuous Case

Remark 8.96 (Contents) This section presents a finite element error analysis of the continuous-in-time discretization of the Smagorinsky model from John and Layton (2002), see also John (2004, Sect. 8.1). The error of a finite element discretization of the Smagorinsky model to the continuous Smagorinsky model is analyzed (and not to the Navier–Stokes equations). In this analysis, a generalization of the Smagorinsky model and pairs of inf-sup stable finite element spaces are considered. In addition, other boundary conditions than no-slip conditions are allowed on a part of the boundary. The goal of the analysis consists in deriving uniform estimates, i.e., with right-hand sides that do not depend on the viscosity ν, with as weak assumptions on the regularity of the solution of the continuous Smagorinsky model as possible. In this respect, the obtained estimates are better than for the Navier–Stokes equations, see Theorem 7.35. Besides the existence and uniqueness of a weak solution of the continuous problem, these estimates are another indicator that the complexity of the Smagorinsky model is smaller than of the Navier–Stokes equations, see Remark 8.22. □

8.3.3.1 The Continuous Problem

Remark 8.97 (The Strong Formulation of the Problem) Let Ω be a bounded domain in $\mathbb{R}^d$, $d \in \{2, 3\}$, with polygonal or polyhedral Lipschitz boundary and suppose that the boundary $\partial\Omega$ is composed of faces $\Gamma_0, \ldots, \Gamma_J$ with $|\Gamma_0| > 0$.

The Smagorinsky model is considered with slip with linear friction and no penetration boundary conditions, see Remark 2.26,

$$\begin{aligned}
\partial_t \boldsymbol{w} - \nabla \cdot ((2\nu + \nu_{\mathrm{T}}) \mathbb{D}(\boldsymbol{w})) & \\
+ (\boldsymbol{w} \cdot \nabla) \boldsymbol{w} + \nabla r &= \boldsymbol{f} \quad \text{in } (0, T] \times \Omega, \\
\nabla \cdot \boldsymbol{w} &= 0 \quad \text{in } (0, T] \times \Omega, \\
\boldsymbol{w} &= \boldsymbol{0} \quad \text{in } (0, T] \times \Gamma_0, \\
\boldsymbol{w} \cdot \boldsymbol{n}_{\partial\Omega} &= 0 \quad \text{in } (0, T] \times \Gamma_j, j = 1, \ldots, J, \\
\boldsymbol{w} \cdot \boldsymbol{\tau}_{j,k} + \beta^{-1} \boldsymbol{n}_{\partial\Omega}^T \mathbb{S}_{\mathrm{sma}}(\boldsymbol{w}, r) \boldsymbol{\tau}_{j,k} &= 0 \quad \text{in } (0, T] \times \Gamma_j, j = 1, \ldots, J, \\
\boldsymbol{w}(0, \cdot) &= \boldsymbol{w}_0 \quad \text{in } \Omega, \\
\int_\Omega r \, d\boldsymbol{x} &= 0 \quad \text{in } (0, T].
\end{aligned} \tag{8.110}$$

Here, $\{\boldsymbol{\tau}_{j,k}\}_{k=1}^{d-1}$ is an orthonormal system of tangential vectors in each point of Γ_j, $1, \ldots, J$, and

$$\mathbb{S}_{\mathrm{sma}}(\boldsymbol{w}, r) = (2\nu + \nu_{\mathrm{T}}) \mathbb{D}(\boldsymbol{w}) - r\mathbb{I}.$$

The turbulent viscosity is given by

$$\nu_{\mathrm{T}} = \nu_0(\delta) + C_S \delta^2 \|\mathbb{D}(\boldsymbol{w})\|_{\mathrm{F}}, \quad \nu_0(\delta) \geq 0,$$

which is a generalization of (8.66). □

Remark 8.98 (The Variational Formulation) Let

$$V = \left\{\boldsymbol{v} \in W^{1,3}(\Omega), \boldsymbol{v} = \boldsymbol{0} \text{ on } \Gamma_0, \boldsymbol{v} \cdot \boldsymbol{n}_{\partial\Omega} = 0 \text{ on } \Gamma_j, j = 1, \ldots, J\right\},$$

$$Q = L_0^2(\Omega).$$

To simplify the notation, whenever $\boldsymbol{\tau}_j$ occurs, it will be understood that the term is summed over the two tangential vectors if $d = 3$, i.e.,

$$\left\| \boldsymbol{v} \cdot \boldsymbol{\tau}_j \right\|^2_{L^2(\Gamma_j)} := \left\| \boldsymbol{v} \cdot \boldsymbol{\tau}_{j,1} \right\|^2_{L^2(\Gamma_j)} + \left\| \boldsymbol{v} \cdot \boldsymbol{\tau}_{j,2} \right\|^2_{L^2(\Gamma_j)} .$$

The variational formulation of (8.110) is derived in the usual way by multiplying the equations with test functions, integrating on Ω, and applying integration by parts. Applying this procedure to the viscous and pressure term gives an integral on the boundary. Using the decomposition

$$\boldsymbol{v} = \sum_{j=1}^{J} (\boldsymbol{v} \cdot \boldsymbol{n}_{\partial\Omega})\, \boldsymbol{n}_{\partial\Omega} + \left(\boldsymbol{v} \cdot \boldsymbol{\tau}_j\right) \boldsymbol{\tau}_j$$

and the definition of the boundary condition gives for this term

$$\begin{aligned}
&\int_{\partial\Omega} -\boldsymbol{n}^T_{\partial\Omega} \left((2\nu + \nu_{\mathrm{T}})\, \mathbb{D}\left(\boldsymbol{w}\right) + r\mathbb{I}\right) \boldsymbol{v}\; d\boldsymbol{s} \\
&= \sum_{j=1}^{J} \int_{\Gamma_j} -\boldsymbol{n}^T_{\partial\Omega} \mathbb{S}_{\mathrm{sma}}\left(\boldsymbol{w}, r\right) \boldsymbol{\tau}_j \left(\boldsymbol{v} \cdot \boldsymbol{\tau}_j\right)\; d\boldsymbol{s} = \sum_{j=1}^{J} \int_{\Gamma_j} \beta \left(\boldsymbol{w} \cdot \boldsymbol{\tau}_j\right) \left(\boldsymbol{v} \cdot \boldsymbol{\tau}_j\right)\; d\boldsymbol{s}.
\end{aligned}$$

The variational problem reads as follows: Find $(\boldsymbol{w}, r) \in V \times Q$ such that for all $t \in (0, T]$ and all $(\boldsymbol{v}, q) \in V \times Q$

$$(\partial_t \boldsymbol{w}, \boldsymbol{v}) + a\left(\boldsymbol{w}, \boldsymbol{w}, \boldsymbol{v}\right) + n_{\mathrm{skew}}\left(\boldsymbol{w}, \boldsymbol{w}, \boldsymbol{v}\right) + (\nabla \cdot \boldsymbol{w}, q) - (\nabla \cdot \boldsymbol{v}, r) = (\boldsymbol{f}, \boldsymbol{v}) \qquad (8.111)$$

with

$$\begin{aligned}
a\left(\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v}\right) = &\left(\left(2\nu + \nu_0\left(\delta\right) + C_S \delta^2 \left\|\mathbb{D}\left(\boldsymbol{u}\right)\right\|_{\mathrm{F}}\right) \mathbb{D}\left(\boldsymbol{w}\right), \mathbb{D}\left(\boldsymbol{v}\right)\right) \\
&+ \mu \left(\nabla \cdot \boldsymbol{w}, \nabla \cdot \boldsymbol{v}\right) + \sum_{j=1}^{J} \beta \left(\boldsymbol{w} \cdot \boldsymbol{\tau}_j, \boldsymbol{v} \cdot \boldsymbol{\tau}_j\right)_{\Gamma_j},
\end{aligned}$$

where $\mu > 0$ is given, and $\boldsymbol{w}\left(0, \boldsymbol{x}\right) = \boldsymbol{w}_0\left(\boldsymbol{x}\right)$. Note that the skew-symmetric form of the nonlinear convective term is for the continuous problem equivalent to the standard convective form, see Remark 6.8, and that the grad-div stabilization vanishes since $\boldsymbol{w}$ is weakly divergence-free. □

Remark 8.99 (Some Tools for the Analysis) Since $\boldsymbol{v} \cdot \boldsymbol{n} = 0$ on $\partial\Omega$ for all $\boldsymbol{v} \in V$, Poincare's inequality (A.12) holds in V. In addition, it is known that Korn's inequality holds in V, i.e.,

$$\|\nabla \boldsymbol{v}\|_{L^3(\Omega)} \le C \|\mathbb{D}\left(\boldsymbol{v}\right)\|_{L^3(\Omega)}. \qquad (8.112)$$

The dual space V' of V is equipped with the norm

$$\|\boldsymbol{\phi}\|_{V'} := \sup_{\boldsymbol{v}\in V\setminus\{\boldsymbol{0}\}} \frac{\displaystyle\int_\Omega \boldsymbol{\phi}\cdot\boldsymbol{v}\,d\boldsymbol{x}}{\|\mathbb{D}(\boldsymbol{v})\|_{L^3(\Omega)}}. \tag{8.113}$$

Note that $\|\mathbb{D}(\boldsymbol{v})\|_{L^3(\Omega)}$ defines a norm in V as a consequence of Poincare's and Korn's inequality. In the case $\partial\Omega = \Gamma_0$, one has $V' = W^{-1,3/2}(\Omega)$ equipped with the norm (8.113). □

Lemma 8.100 (Strong Monotonicity of the Nonlinear Viscous Term) *There is a constant* $\underline{C}$ *such that for all* $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in W^{1,3}(\Omega)$

$$\left(\|\mathbb{D}(\boldsymbol{u})\|_{\mathrm{F}}\mathbb{D}(\boldsymbol{u}) - \|\mathbb{D}(\boldsymbol{v})\|_{\mathrm{F}}\mathbb{D}(\boldsymbol{v}), \mathbb{D}(\boldsymbol{u}-\boldsymbol{v})\right) \ge \underline{C}\,\|\mathbb{D}(\boldsymbol{u}-\boldsymbol{v})\|_{L^3(\Omega)}^3, \tag{8.114}$$

i.e., the strong monotonicity holds.

Proof The proof of Lemma 8.88 can be performed also for the deformation tensor instead of the gradient. All partial derivatives have to be replaced by the respective entries of the deformation tensor, e.g., $\frac{\partial w_i'}{\partial x_j}$ by $\mathbb{D}_{ij}(\boldsymbol{w}')$. ■

Lemma 8.101 (Norm Equivalence for a Tensor) *Let* $A\in L^3(\Omega)$ *with* $A(\boldsymbol{x})\in\mathbb{R}^{d\times d}$ *for every* $\boldsymbol{x}\in\Omega$*, then*

$$\|A\|_{L^3(\Omega)} \le \|\,\|A\|_{\mathrm{F}}\|_{L^3(\Omega)} \le C(d)\,\|A\|_{L^3(\Omega)}. \tag{8.115}$$

Proof It is

$$\|A\|_{L^3(\Omega)}^3 = \int_\Omega \sum_{i,j=1}^d \left|a_{ij}\right|^3\,d\boldsymbol{x}, \quad \|\,\|A\|_{\mathrm{F}}\|_{L^3(\Omega)}^3 = \int_\Omega \left(\sum_{i,j=1}^d a_{ij}^2\right)^{3/2} d\boldsymbol{x}.$$

Since all matrix norms are equivalent, there are constants $0 < C_1(d) < C_2(d)$ such that

$$C_1(d)\left(\sum_{i,j=1}^d \left|a_{ij}\right|^3\right)^{1/3} \le \left(\sum_{i,j=1}^d a_{ij}^2\right)^{1/2} \le C_2(d)\left(\sum_{i,j=1}^d \left|a_{ij}\right|^3\right)^{1/3}. \tag{8.116}$$

From (A.4), it follows with $p = 3/2$ that one can choose $C_1(d) = 1$. Raising (8.116) to the power 3 and integrating on Ω proves (8.115). ■

Lemma 8.102 (Local Lipschitz Continuity of the Nonlinear Viscous Term) *There is a constant* $\overline{C}$ *such that for all* $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in W^{1,3}(\Omega)$

$$\begin{aligned}&\left(\|\mathbb{D}(\boldsymbol{u})\|_{\mathrm{F}}\mathbb{D}(\boldsymbol{u}) - \|\mathbb{D}(\boldsymbol{v})\|_{\mathrm{F}}\mathbb{D}(\boldsymbol{v}), \mathbb{D}(\boldsymbol{w})\right) \\ &\le \overline{C}C_L\,\|\mathbb{D}(\boldsymbol{u}-\boldsymbol{v})\|_{L^3(\Omega)}\,\|\mathbb{D}(\boldsymbol{w})\|_{L^3(\Omega)}\end{aligned} \tag{8.117}$$

with $C_L = \max\left\{\|\mathbb{D}(\boldsymbol{u})\|_{L^3(\Omega)}, \|\mathbb{D}(\boldsymbol{v})\|_{L^3(\Omega)}\right\}$, *which is the so-called local Lipschitz continuity.*

Proof Applying Hölder's inequality (A.9) gives

$$\begin{aligned}
&\left(\|\mathbb{D}(\boldsymbol{u})\|_{\mathrm{F}}\mathbb{D}(\boldsymbol{u}) - \|\mathbb{D}(\boldsymbol{v})\|_{\mathrm{F}}\mathbb{D}(\boldsymbol{v}), \mathbb{D}(\boldsymbol{w})\right)\\
&= \left(\|\mathbb{D}(\boldsymbol{u})\|_{\mathrm{F}}\mathbb{D}(\boldsymbol{u}) - \|\mathbb{D}(\boldsymbol{u})\|_{\mathrm{F}}\mathbb{D}(\boldsymbol{v}), \mathbb{D}(\boldsymbol{w})\right) + \left(\|\mathbb{D}(\boldsymbol{u})\|_{\mathrm{F}}\mathbb{D}(\boldsymbol{v}) - \|\mathbb{D}(\boldsymbol{v})\|_{\mathrm{F}}\mathbb{D}(\boldsymbol{v}), \mathbb{D}(\boldsymbol{w})\right)\\
&\le \left\|\|\mathbb{D}(\boldsymbol{u})\|_{\mathrm{F}}\right\|_{L^3(\Omega)}\|\mathbb{D}(\boldsymbol{u}-\boldsymbol{v})\|_{L^3(\Omega)}\|\mathbb{D}(\boldsymbol{w})\|_{L^3(\Omega)}\\
&\quad + \left\|\|\mathbb{D}(\boldsymbol{u})\|_{\mathrm{F}} - \|\mathbb{D}(\boldsymbol{v})\|_{\mathrm{F}}\right\|_{L^3(\Omega)}\|\mathbb{D}(\boldsymbol{v})\|_{L^3(\Omega)}\|\mathbb{D}(\boldsymbol{w})\|_{L^3(\Omega)}.
\end{aligned} \tag{8.118}$$

Using the triangle inequality and (8.115) yields

$$\left\|\|\mathbb{D}(\boldsymbol{u})\|_{\mathrm{F}} - \|\mathbb{D}(\boldsymbol{v})\|_{\mathrm{F}}\right\|_{L^3(\Omega)} \le \left\|\|\mathbb{D}(\boldsymbol{u}) - \mathbb{D}(\boldsymbol{v})\|_{\mathrm{F}}\right\|_{L^3(\Omega)} \le C\|\mathbb{D}(\boldsymbol{u}-\boldsymbol{v})\|_{L^3(\Omega)}.$$

Inserting this estimate in (8.118) proves the local Lipschitz continuity. ■

Lemma 8.103 (Energy Inequality for *w*) *Any solution of* (8.111) *satisfies*

$$\begin{aligned}
\frac{1}{2}\|\boldsymbol{w}(T)\|_{L^2(\Omega)}^2 &+ \int_0^T\left(\sum_{j=1}^J \beta\left\|\boldsymbol{w}\cdot\boldsymbol{\tau}_j\right\|_{L^2(\Gamma_j)}^2 + (2\nu + \nu_0(\delta))\|\mathbb{D}(\boldsymbol{w})\|_{L^2(\Omega)}^2\right.\\
&\left. + \underline{C}C_S\delta^2\|\mathbb{D}(\boldsymbol{w})\|_{L^3(\Omega)}^3\right)dt \le \frac{1}{2}\|\boldsymbol{w}_0\|_{L^2(\Omega)}^2 + \int_0^T(\boldsymbol{f},\boldsymbol{w})\ dt.
\end{aligned} \tag{8.119}$$

Proof Choosing $(\boldsymbol{v}, q) = (\boldsymbol{w}, r)$ in (8.111) gives

$$(\partial_t\boldsymbol{w}, \boldsymbol{w}) + a(\boldsymbol{w}, \boldsymbol{w}, \boldsymbol{w}) + n_{\mathrm{skew}}(\boldsymbol{w}, \boldsymbol{w}, \boldsymbol{w}) + (\nabla\cdot\boldsymbol{w}, r) - (\nabla\cdot\boldsymbol{w}, r) = (\boldsymbol{f}, \boldsymbol{w}).$$

The skew symmetric nonlinear convective term vanishes, see (6.26).

Since every $q \in L^2(\Omega)$ admits a decomposition $q = q_0 + C$ with $q_0 \in Q$ and C is a constant, see Remark 4.70, it follows that

$$\begin{aligned}
(\nabla\cdot\boldsymbol{v}, q) &= (\nabla\cdot\boldsymbol{v}, q_0 + C) = (\nabla\cdot\boldsymbol{v}, C)\\
&= -\int_\Gamma C\boldsymbol{v}\cdot\boldsymbol{n}_{\partial\Omega}\ d\boldsymbol{s} - (\boldsymbol{v}, \nabla C) = 0 \quad \forall\ \boldsymbol{v}\in V_{\mathrm{div}}, q\in L^2(\Omega),
\end{aligned} \tag{8.120}$$

because $\boldsymbol{v}\cdot\boldsymbol{n}_{\partial\Omega} = 0$ on the whole boundary. Thus, $\mu(\nabla\cdot\boldsymbol{w}, \nabla\cdot\boldsymbol{w})$ vanishes because $\boldsymbol{w}\in V_{\mathrm{div}}$, $\nabla\cdot\boldsymbol{w}\in L^2(\Omega)$, and (8.120).

In addition, it follows from (8.114) that

$$\left(\|\mathbb{D}(\boldsymbol{w})\|_{\mathrm{F}}\mathbb{D}(\boldsymbol{w}), \mathbb{D}(\boldsymbol{v})\right) \ge \underline{C}\|\mathbb{D}(\boldsymbol{w})\|_{L^3(\Omega)}^3$$

such that

$$(\partial_t \boldsymbol{w}, \boldsymbol{w}) + (2\nu + \nu_0(\delta)) \|\mathbb{D}(\boldsymbol{w})\|^2_{L^2(\Omega)} + \underline{C} C_S \delta^2 \|\mathbb{D}(\boldsymbol{w})\|^3_{L^3(\Omega)} + \sum_{j=1}^{J} \beta \left\|\boldsymbol{w} \cdot \boldsymbol{\tau}_j\right\|^2_{L^2(\Gamma_j)} \le (\boldsymbol{f}, \boldsymbol{w}).$$

Integration on $(0,T)$ gives the statement of the lemma. ■

Lemma 8.104 (Stability Estimates for $\boldsymbol{w}$ Uniformly in ν) *The velocity component* $\boldsymbol{w}$ *of the solution of* (8.111) *satisfies for* $T > 0$

$$\frac{1}{2}\|\boldsymbol{w}(T)\|^2_{L^2(\Omega)} + \int_0^T \left(\sum_{j=1}^{J} \beta \left\|\boldsymbol{w} \cdot \boldsymbol{\tau}_j\right\|^2_{L^2(\Gamma_j)} + (2\nu + \nu_0(\delta)) \|\mathbb{D}(\boldsymbol{w})\|^2_{L^2(\Omega)} + \frac{2}{3}\underline{C} C_S \delta^2 \|\mathbb{D}(\boldsymbol{w})\|^3_{L^3(\Omega)} \right) dt \tag{8.121}$$

$$\le \frac{1}{2}\|\boldsymbol{w}_0\|^2_{L^2(\Omega)} + \frac{2}{3}\left(\underline{C} C_S\right)^{-1/2} \delta^{-1} \left(\sup_{\boldsymbol{v} \in L^3(0,T;V)} \frac{\int_0^T (\boldsymbol{f}, \boldsymbol{v})\, dt}{\|\mathbb{D}(\boldsymbol{v})\|_{L^3(0,T;L^3(\Omega))}} \right)^{3/2}$$

and

$$\frac{1}{2}\|\boldsymbol{w}(T)\|^2_{L^2(\Omega)} + \int_0^T e^{T-t} \left(\sum_{j=1}^{J} \beta \left\|\boldsymbol{w} \cdot \boldsymbol{\tau}_j\right\|^2_{L^2(\Gamma_j)} + (2\nu + \nu_0(\delta)) \|\mathbb{D}(\boldsymbol{w})\|^2_{L^2(\Omega)} + \underline{C} C_S \delta^2 \|\mathbb{D}(\boldsymbol{w})\|^3_{L^3(\Omega)} \right) dt$$

$$\le \frac{e^T}{2}\|\boldsymbol{w}_0\|^2_{L^2(\Omega)} + \frac{1}{2}\int_0^T e^{T-t} \|\boldsymbol{f}\|^2_{L^2(\Omega)}\, dt. \tag{8.122}$$

Proof Young's inequality (A.5) and the first stability estimate (8.119). It is

$$\int_0^T (\boldsymbol{f}, \boldsymbol{w})\, dt = \|\mathbb{D}(\boldsymbol{w})\|_{L^3(0,T;L^3(\Omega))} \frac{\int_0^T (\boldsymbol{f}, \boldsymbol{w})\, dt}{\|\mathbb{D}(\boldsymbol{w})\|_{L^3(0,T;L^3(\Omega))}}$$

$$\le \|\mathbb{D}(\boldsymbol{w})\|_{L^3(0,T;L^3(\Omega))} \sup_{\boldsymbol{v} \in L^3(0,T;L^3(\Omega))} \frac{\int_0^T (\boldsymbol{f}, \boldsymbol{v})\, dt}{\|\mathbb{D}(\boldsymbol{v})\|_{L^3(0,T;L^3(\Omega))}}$$

$$\le \frac{\underline{C} C_S \delta^2}{3} \|\mathbb{D}(\boldsymbol{w})\|^3_{L^3(0,T;L^3(\Omega))} + \frac{2}{3}\left(\underline{C} C_S\right)^{-1/2} \delta^{-1} \left(\sup_{\boldsymbol{v} \in L^3(0,T;L^3(\Omega))} \frac{\int_0^T (\boldsymbol{f}, \boldsymbol{v})\, dt}{\|\mathbb{D}(\boldsymbol{v})\|_{L^3(0,T;L^3(\Omega))}} \right)^{3/2}.$$

Absorbing the first term on the right-hand side in (8.119) gives (8.121).

Equation (8.122): In the same way as in the proof of Lemma 8.103, one obtains with the Cauchy–Schwarz inequality (A.10) and Young's inequality

$$\frac{1}{2}\frac{d}{dt}\|\boldsymbol{w}\|_{L^2(\Omega)}^2 + (2\nu + \nu_0(\delta))\|\mathbb{D}(\boldsymbol{w})\|_{L^2(\Omega)}^2$$
$$+\underline{C}C_S\delta^2\|\mathbb{D}(\boldsymbol{w})\|_{L^3(\Omega)}^3 + \sum_{j=1}^{J}\beta\|\boldsymbol{w}\cdot\boldsymbol{\tau}_j\|_{L^2(\Gamma_j)}^2 \le \frac{\|\boldsymbol{f}\|_{L^2(\Omega)}^2}{2} + \frac{\|\boldsymbol{w}\|_{L^2(\Omega)}^2}{2}.$$

Multiplying this equality with e^{T-t}, integrating on $(0,T)$, and applying integration by parts yields

$$\int_0^T e^{T-t}\left(\frac{d}{dt}\|\boldsymbol{w}\|_{L^2(\Omega)}^2 - \|\boldsymbol{w}\|_{L^2(\Omega)}^2\right)dt$$
$$= e^{T-t}\|\boldsymbol{w}\|_{L^2(\Omega)}^2\Big|_{t=0}^{t=T} + \int_0^T e^{T-t}\|\boldsymbol{w}\|_{L^2(\Omega)}^2\,dt - \int_0^T e^{T-t}\|\boldsymbol{w}\|_{L^2(\Omega)}^2\,dt$$
$$= \|\boldsymbol{w}(T)\|_{L^2(\Omega)}^2 - e^T\|\boldsymbol{w}(0)\|_{L^2(\Omega)}^2.$$

Using this equality, (8.122) is proved. ■

Remark 8.105 (On the Stability Estimate (8.121)) The right-hand side of the stability estimate (8.121) is independent of ν but it depends on inverse powers of δ. Thus, if $\delta \to 0$, the right-hand side blows up. This behavior is the natural one since otherwise one would find in the limit a uniform stability estimate for the solution of the Navier–Stokes equations. □

Lemma 8.106 (Stability of $\partial_t \boldsymbol{w}$ Uniformly in ν) *Let $(\boldsymbol{w}, r)$ be a solution of (8.111). Then, there is a constant C independent of ν such that for almost all $t \in [0,T]$*

$$\|\partial_t\boldsymbol{w}\|_{V'} \le C\Big(\|\boldsymbol{w}\|_{L^3(\Omega)}^2 + \|r\|_{L^{3/2}(\Omega)} + (2\nu + \nu_0(\delta))\|\mathbb{D}(\boldsymbol{w})\|_{L^{3/2}(\Omega)}$$
$$+C_S\delta^2\|\mathbb{D}(\boldsymbol{w})\|_{L^3(\Omega)}^2 + \|\boldsymbol{f}\|_{V'}\Big) \tag{8.123}$$

and

$$\|\partial_t\boldsymbol{w}\|_{L^{3/2}(0,T;V')}^{3/2} \le C\Big(\|\boldsymbol{w}\|_{L^3(0,T;L^3(\Omega))}^3 + \|r\|_{L^{3/2}(0,T;L^{3/2}(\Omega))}^{3/2}$$
$$+ (2\nu + \nu_0(\delta))\|\mathbb{D}(\boldsymbol{w})\|_{L^{3/2}(0,T;L^{3/2}(\Omega))}^{3/2}$$
$$+C_S\delta^2\|\mathbb{D}(\boldsymbol{w})\|_{L^3(0,T;L^3(\Omega))}^3 + \|\boldsymbol{f}\|_{L^{3/2}(0,T;V')}^{3/2}\Big).$$

Proof Multiplying (8.110) with $\boldsymbol{v}$, integrating over Ω, dividing the resulting equation by $\|\mathbb{D}(\boldsymbol{v})\|_{L^3(\Omega)}$, taking the supremum over $\boldsymbol{v} \in V$, and applying the triangle inequality gives

$$\begin{aligned}\|\partial_t \boldsymbol{w}\|_{V'} \le \left\|\nabla \cdot \left(\boldsymbol{w}\boldsymbol{w}^T\right)\right\|_{V'} + \|\nabla r\|_{V'} + C_S\delta^2 \left\|\nabla \cdot \left(\|\mathbb{D}(\boldsymbol{w})\|_{\mathrm{F}}\, \mathbb{D}(\boldsymbol{w})\right)\right\|_{V'} \\ + (2\nu + \nu_0(\delta)) \|\nabla \cdot \mathbb{D}(\boldsymbol{w})\|_{V'} + \|\boldsymbol{f}\|_{V'}.\end{aligned}$$

The definition of the norm (8.113), integration by parts, using $\boldsymbol{v} \cdot \boldsymbol{n}_{\partial\Omega} = 0$ on $\partial\Omega$ for $\boldsymbol{v} \in V$, Hölder's inequality (A.9), estimating the norm of the divergence by the norm of the gradient, and Korn's inequality (8.112) leads to, e.g.,

$$\begin{aligned}\|\nabla r\|_{V'} &= \sup_{\boldsymbol{v}\in V\setminus\{\boldsymbol{0}\}} \frac{-(\nabla\cdot\boldsymbol{v}, r)}{\|\mathbb{D}(\boldsymbol{v})\|_{L^3(\Omega)}} \le \sup_{\boldsymbol{v}\in V\setminus\{\boldsymbol{0}\}} \frac{\|r\|_{L^{3/2}(\Omega)} \|\nabla\cdot\boldsymbol{v}\|_{L^3(\Omega)}}{\|\mathbb{D}(\boldsymbol{v})\|_{L^3(\Omega)}} \\ &\le C \sup_{\boldsymbol{v}\in V\setminus\{\boldsymbol{0}\}} \frac{\|r\|_{L^{3/2}(\Omega)} \|\nabla\boldsymbol{v}\|_{L^3(\Omega)}}{\|\mathbb{D}(\boldsymbol{v})\|_{L^3(\Omega)}} \le C\|r\|_{L^{3/2}(\Omega)}.\end{aligned}$$

Note that an estimate of the norm of the divergence by the norm of the gradient of form (3.40) can be proved for any $L^p(\Omega)$ norm by using Hölder's inequality instead of the Cauchy–Schwarz inequality in the proof of (3.40).

The other terms are estimated in the same way, using the following estimates. With the Cauchy–Schwarz (A.10) inequality, it is

$$\begin{aligned}\left\|\boldsymbol{w}\boldsymbol{w}^T\right\|_{L^{3/2}(\Omega)} &= \left(\int_\Omega \sum_{i,j=1}^d \left|w_i w_j\right|^{3/2}\, d\boldsymbol{x}\right)^{2/3} \\ &\le \left(\int_\Omega \left(\sum_{i=1}^d |w_i|^3\right)^{1/2} \left(\sum_{i=j}^d \left|w_j\right|^3\right)^{1/2} d\boldsymbol{x}\right)^{2/3} \\ &= \left(\int_\Omega \sum_{i=1}^d |w_i|^3\, d\boldsymbol{x}\right)^{2/3} = \|\boldsymbol{w}\|^2_{L^3(\Omega)}.\end{aligned}$$

Using Hölder's inequality and (8.115), one finds that

$$\begin{aligned}\left\|\|\mathbb{D}(\boldsymbol{w})\|_{\mathrm{F}}\,\mathbb{D}(\boldsymbol{w})\right\|_{L^{3/2}(\Omega)} &= \left(\int_\Omega \sum_{i,j=1}^d \left|\|\mathbb{D}(\boldsymbol{w})\|_{\mathrm{F}}\,\mathbb{D}_{ij}(\boldsymbol{w})\right|^{3/2}\, d\boldsymbol{x}\right)^{2/3} \\ &\le C\left(\int_\Omega \|\mathbb{D}(\boldsymbol{w})\|_{\mathrm{F}}^{3/2} \left(\sum_{i,j=1}^d \left|\mathbb{D}_{ij}(\boldsymbol{w})\right|^2\right)^{3/4} d\boldsymbol{x}\right)^{2/3}\end{aligned}$$

$$= C \left(\int_{\Omega} \|\mathbb{D}(\boldsymbol{w})\|_{\mathrm{F}}^{3/2} \|\mathbb{D}(\boldsymbol{w})\|_{\mathrm{F}}^{3/2} \ d\boldsymbol{x} \right)^{2/3}$$

$$= C \left\| \|\mathbb{D}(\boldsymbol{w})\|_{\mathrm{F}} \right\|_{L^3(\Omega)}^2 \le C \|\mathbb{D}(\boldsymbol{w})\|_{L^3(\Omega)}^2 .$$

The second inequality stated in the lemma follows by raising both sides of (8.123) to the power $3/2$, integrating in time, and absorbing some powers of coefficients and parameters in the constant. ■

8.3.3.2 The Finite Element Problem

Remark 8.107 (Finite Element Spaces) The discrete problem is defined in finite element spaces V^h and Q^h with $V^h \subset V$ and $Q^h \subset Q$. It will be assumed that the spaces V^h and Q^h satisfy the discrete inf-sup condition

$$\inf_{\substack{q^h \in Q^h \\ q^h \neq 0}} \sup_{\substack{\boldsymbol{v}^h \in V^h \\ \boldsymbol{v}^h \neq \boldsymbol{0}}} \frac{\left(\nabla \cdot \boldsymbol{v}^h, q^h\right)}{\left\|q^h\right\|_{L^2(\Omega)} \left\|\boldsymbol{v}^h\right\|_V} \ge \beta_{\mathrm{is,sma}}^h > 0,$$

where $\beta_{\mathrm{is,sma}}$ is independent of h and

$$\left\|\boldsymbol{v}^h\right\|_V^2 = \left\|\mathbb{D}\left(\boldsymbol{v}^h\right)\right\|_{L^2(\Omega)}^2 + \sum_{j=1}^{J} \left\|\boldsymbol{v}^h \cdot \boldsymbol{\tau}_j\right\|_{H^{1/2}(\Gamma_j)}^2 .$$

□

Remark 8.108 (The Continuous-in-Time Finite Element Problem) The continuous-in-time finite element problem reads as follows: Find $\left(\boldsymbol{w}^h, r^h\right) \in V^h \times Q^h$ such that

$$\begin{aligned} \left(\partial_t \boldsymbol{w}^h, \boldsymbol{v}^h\right) + a\left(\boldsymbol{w}^h, \boldsymbol{w}^h, \boldsymbol{v}^h\right) + n_{\mathrm{skew}}\left(\boldsymbol{w}^h, \boldsymbol{w}^h, \boldsymbol{v}^h\right) \\ + \left(\nabla \cdot \boldsymbol{w}^h, q^h\right) - \left(\nabla \cdot \boldsymbol{v}^h, r^h\right) = \left(\boldsymbol{f}, \boldsymbol{v}^h\right) \end{aligned} \tag{8.124}$$

for all $\left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h$ where $\boldsymbol{w}^h(0, \boldsymbol{x})$ is an approximation to $\boldsymbol{w}_0(\boldsymbol{x})$. □

Lemma 8.109 (Energy Inequality for the Finite Element Solution) *A solution of* (8.124) *satisfies*

$$\begin{aligned} &\frac{1}{2}\left\|\boldsymbol{w}^h(T)\right\|_{L^2(\Omega)}^2 + \int_0^T \left(\mu \left\|\nabla \cdot \boldsymbol{w}^h\right\|_{L^2(\Omega)}^2 + \sum_{j=1}^{J} \beta \left\|\boldsymbol{w}^h \cdot \boldsymbol{\tau}_j\right\|_{L^2(\Gamma_j)}^2 \right. \\ &\quad \left. + \left(2\nu + \nu_0(\delta)\right) \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^2(\Omega)}^2 + \underline{C} C_S \delta^2 \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^3 \right) dt \\ &\le \frac{1}{2}\left\|\boldsymbol{w}_0\right\|_{L^2(\Omega)}^2 + \int_0^T \left(\boldsymbol{f}, \boldsymbol{w}^h\right) \ dt. \end{aligned}$$

Proof The proof proceeds in the same way as the proof of Lemma 8.103. ■

Theorem 8.110 (Stability Estimates for w^h Uniformly in ν) *The velocity component $\boldsymbol{w}^h$ of the solution of* (8.124) *satisfies for* $T > 0$

$$\begin{aligned} &\frac{1}{2}\left\|\boldsymbol{w}^h(T)\right\|^2_{L^2(\Omega)} + \int_0^T \Bigg(\mu\left\|\nabla\cdot\boldsymbol{w}^h\right\|^2_{L^2(\Omega)} + \sum_{j=1}^J \beta\left\|\boldsymbol{w}^h\cdot\boldsymbol{\tau}_j\right\|^2_{L^2(\Gamma_j)} \\ &\quad + (2\nu + \nu_0(\delta))\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|^2_{L^2(\Omega)} + \frac{2}{3}\underline{C}C_S\delta^2\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|^3_{L^3(\Omega)}\Bigg)\,dt \qquad (8.125) \\ &\le \frac{1}{2}\left\|\boldsymbol{w}_0^h\right\|^2_{L^2(\Omega)} + \frac{2}{3}\left(\underline{C}C_S\right)^{-1/2}\delta^{-1}\left(\sup_{\boldsymbol{v}\in L^3(0,T;V)}\frac{\int_0^T(\boldsymbol{f},\boldsymbol{v})\,dt}{\|\mathbb{D}(\boldsymbol{v})\|_{L^3(0,T;L^3(\Omega))}}\right)^{3/2} \end{aligned}$$

and

$$\begin{aligned} &\frac{1}{2}\left\|\boldsymbol{w}^h(T)\right\|^2_{L^2(\Omega)} + \int_0^T e^{T-t}\Bigg(\mu\left\|\nabla\cdot\boldsymbol{w}^h\right\|^2_{L^2(\Omega)} + \sum_{j=1}^J \beta\left\|\boldsymbol{w}^h\cdot\boldsymbol{\tau}_j\right\|^2_{L^2(\Gamma_j)} \\ &\quad + (2\nu + \nu_0(\delta))\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|^2_{L^2(\Omega)} + \underline{C}C_S\delta^2\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|^3_{L^3(\Omega)}\Bigg)\,dt \\ &\le \frac{e^T}{2}\left\|\boldsymbol{w}_0^h\right\|^2_{L^2(\Omega)} + \frac{1}{2}\int_0^T e^{T-t}\|\boldsymbol{f}\|^2_{L^2(\Omega)}\,dt. \qquad (8.126) \end{aligned}$$

Proof The proof is performed in the same way as the proof of Lemma 8.104. ■

Remark 8.111 (Goal and Summary of the Finite Element Error Analysis) The goal of the analysis is to estimate the error $\left\|\boldsymbol{w}-\boldsymbol{w}^h\right\|$ in appropriate norms under consideration of the following aspects:

- The error bound should be independent of ν.
- The assumption on the regularity of the solution of the variational problem (8.111) should be as weak as possible.

The following results will be presented.

- First, the natural regularity

$$\nabla\boldsymbol{w} \in L^3\left(0,T;L^3(\Omega)\right), \qquad (8.127)$$

  for the variational formulation of the Smagorinsky model will be assumed. This regularity ensures that the Smagorinsky term is well defined, see (8.72). A finite element error analysis, using a standard approach sketched in Remark 8.112, can be performed for the case $\nu_0(\delta) > 0$. The error estimate is given in Theorem 8.118.

- Second, the case $\nu_0(\delta) = 0$ is discussed for the regularity assumption (8.127), see Remark 8.119. It will be shown that Gronwall's lemma cannot be applied. Thus, a uniform estimate for this case is open.
- Finally, the situation of assuming higher regularity of $\boldsymbol{w}$ uniformly in ν and $\nu_0(\delta) \geq 0$ is studied and a uniform estimate is presented in Theorem 8.120.

□

Remark 8.112 (Outline of the Proof) Since the error analysis involves a lot of technical details, a plan of the proof is given in advance, compare also the proof of Theorem 7.35.

1. The variational formulation (8.111) and the discrete problem (8.124) with arbitrary test functions $\left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h$ are subtracted to derive the error Eq. (8.131).
2. One chooses an arbitrary $\tilde{\boldsymbol{w}}^h \in V_{\mathrm{div}}^h$ and splits the error $\boldsymbol{e} = \left(\boldsymbol{w} - \tilde{\boldsymbol{w}}^h\right) - \left(\boldsymbol{w}^h - \tilde{\boldsymbol{w}}^h\right) = \boldsymbol{\eta} - \boldsymbol{\phi}^h$. Then, one takes $\boldsymbol{\phi}^h$ as test function in (8.131) to derive (8.132).
3. The left-hand side of the error Eq. (8.132) is bounded from below by the strong monotonicity (8.114) and the right-hand side of (8.132) from above by the local Lipschitz continuity (8.117) of the Smagorinsky term.
4. Estimating the other terms, in particular the nonlinear convective terms, one derives a differential inequality of the form

$$\frac{d}{dt}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 + \tilde{g}\left(\boldsymbol{\phi}^h\right) \leq g(\boldsymbol{\eta}) + \lambda\left(t, \boldsymbol{w}, \boldsymbol{w}^h\right)\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 \tag{8.128}$$

   with $0 \leq \tilde{g}\left(\boldsymbol{\phi}^h\right), g(\boldsymbol{\eta}), \lambda\left(t, \boldsymbol{w}, \boldsymbol{w}^h\right)$ and $\tilde{g}\left(\boldsymbol{\phi}^h\right) \in L^1(0, T)$.
5. The term $g(\boldsymbol{\eta})$ is bounded from above uniformly in ν by the stability estimates already proved in this section. The bounds have to be in $L^1(0, T)$ uniformly in ν.
6. One shows that $\lambda\left(t, \boldsymbol{w}, \boldsymbol{w}^h\right) \in L^1(0, T)$ uniformly in ν by the stability estimates.
7. The application of Gronwall's lemma, Lemma A.55, to (8.128) gives an error estimate for $\boldsymbol{\phi}^h$.
8. The error estimate for $\boldsymbol{w} - \boldsymbol{w}^h$ is obtained by the triangle inequality. The terms containing $\boldsymbol{\eta}$ are bounded independently of ν by best approximation error estimates of the finite element spaces.

□

Lemma 8.113 (Differential Inequality for Performing the Error Estimates) *Let $\boldsymbol{e} = \boldsymbol{w} - \boldsymbol{w}^h$ denote the error, let $\tilde{\boldsymbol{w}}^h \in V_{\mathrm{div}}^h$ be arbitrary, and consider the decomposition*

$$\boldsymbol{e} = \left(\boldsymbol{w} - \tilde{\boldsymbol{w}}^h\right) - \left(\boldsymbol{w}^h - \tilde{\boldsymbol{w}}^h\right) = \boldsymbol{\eta} - \boldsymbol{\phi}^h. \tag{8.129}$$

If $\mu > 0$*, then it holds the estimate*

$$
\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 + \frac{\mu}{2}\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 + \frac{2\nu+\nu_0(\delta)}{2}\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)}^2 \\
&\quad + \frac{\underline{C}C_S\delta^2}{3}\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)}^3 + \sum_{j=1}^{J}\frac{\beta}{2}\left\|\boldsymbol{\phi}^h\cdot\boldsymbol{\tau}_j\right\|_{L^2(\Gamma_j)}^2 \\
&\le \frac{2}{3\left(\underline{C}C_S\right)^{1/2}\delta}\left\|\partial_t\boldsymbol{\eta}\right\|_{V'}^{3/2} + \mu\left\|\nabla\cdot\boldsymbol{\eta}\right\|_{L^2(\Omega)}^2 + \frac{2\nu+\nu_0(\delta)}{2}\left\|\mathbb{D}(\boldsymbol{\eta})\right\|_{L^2(\Omega)}^2 \\
&\quad + \frac{2\overline{C}^{3/2}C_L^{3/2}C_S\delta^2}{3\underline{C}^{1/2}}\left\|\mathbb{D}(\boldsymbol{\eta})\right\|_{L^3(\Omega)}^{3/2} + \sum_{j=1}^{J}\frac{\beta}{2}\left\|\boldsymbol{\eta}\cdot\boldsymbol{\tau}_j\right\|_{L^2(\Gamma_j)}^2 \qquad (8.130) \\
&\quad + \left|n_{\text{skew}}\left(\boldsymbol{w},\boldsymbol{w},\boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\boldsymbol{w}^h,\boldsymbol{w}^h,\boldsymbol{\phi}^h\right)\right| + \frac{1}{\mu}\left\|r-q^h\right\|_{L^2(\Omega)}^2 .
\end{aligned}
$$

Proof By the choice of $\tilde{\boldsymbol{w}}^h$, it follows that $\boldsymbol{\phi}^h \in V_{\text{div}}^h$. Subtracting (8.124) from (8.111) gives for all $\boldsymbol{v}^h \in V_{\text{div}}^h$ and $q^h \in Q_h$

$$
\begin{aligned}
&\left(\partial_t\boldsymbol{e},\boldsymbol{v}^h\right) + a\left(\boldsymbol{w},\boldsymbol{w},\boldsymbol{v}^h\right) - a\left(\boldsymbol{w}^h,\boldsymbol{w}^h,\boldsymbol{v}^h\right) + n_{\text{skew}}\left(\boldsymbol{w},\boldsymbol{w},\boldsymbol{v}_h\right) \qquad (8.131) \\
&\qquad - n_{\text{skew}}\left(\boldsymbol{w}^h,\boldsymbol{w}^h,\boldsymbol{v}_h\right) - \left(r-q^h,\nabla\cdot\boldsymbol{v}^h\right) = 0.
\end{aligned}
$$

One obtains for $\boldsymbol{v}^h = \boldsymbol{\phi}^h$

$$
\begin{aligned}
&\left(\partial_t\boldsymbol{\phi}^h,\boldsymbol{\phi}^h\right) + a\left(\boldsymbol{w}^h,\boldsymbol{w}^h,\boldsymbol{\phi}^h\right) - a\left(\tilde{\boldsymbol{w}}^h,\tilde{\boldsymbol{w}}^h,\boldsymbol{\phi}^h\right) \qquad (8.132) \\
&= \left(\partial_t\boldsymbol{\eta},\boldsymbol{\phi}^h\right) + a\left(\boldsymbol{w},\boldsymbol{w},\boldsymbol{\phi}^h\right) - a\left(\tilde{\boldsymbol{w}}^h,\tilde{\boldsymbol{w}}^h,\boldsymbol{\phi}^h\right) + n_{\text{skew}}\left(\boldsymbol{w},\boldsymbol{w},\boldsymbol{\phi}^h\right) \\
&\quad - n_{\text{skew}}\left(\boldsymbol{w}^h,\boldsymbol{w}^h,\boldsymbol{\phi}^h\right) - \left(r-q^h,\nabla\cdot\boldsymbol{\phi}^h\right).
\end{aligned}
$$

The monotonicity of $a(\cdot,\cdot,\cdot)$, (8.114), implies

$$
\begin{aligned}
&a\left(\boldsymbol{w}^h,\boldsymbol{w}^h,\boldsymbol{\phi}^h\right) - a\left(\tilde{\boldsymbol{w}}^h,\tilde{\boldsymbol{w}}^h,\boldsymbol{\phi}^h\right) \\
&\ge \mu\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 + (2\nu+\nu_0(\delta))\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)}^2 + \underline{C}C_S\delta^2\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)}^3 \\
&\quad + \sum_{j=1}^{J}\beta\left\|\boldsymbol{\phi}^h\cdot\boldsymbol{\tau}_j\right\|_{L^2(\Gamma_j)}^2 .
\end{aligned}
$$

For bounding the right-hand side of (8.132), one has to get norms of $\boldsymbol{\phi}^h$ that can be absorbed from the left-hand side or $\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2$ for the application of Gronwall's lemma, and the norms of $\boldsymbol{\eta}$ have to be bounded by the stability estimates given at the beginning of this section. The local Lipschitz continuity of the trilinear form, (8.117), gives the estimate

$$
\begin{aligned}
& a\left(\boldsymbol{w}, \boldsymbol{w}, \boldsymbol{\phi}^h\right) - a\left(\tilde{\boldsymbol{w}}^h, \tilde{\boldsymbol{w}}^h, \boldsymbol{\phi}^h\right) \\
& \leq \mu \left\|\nabla \cdot \boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \left\|\nabla \cdot \boldsymbol{\eta}\right\|_{L^2(\Omega)} + (2\nu + \nu_0(\delta)) \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)} \left\|\mathbb{D}(\boldsymbol{\eta})\right\|_{L^2(\Omega)} \\
& \quad + \overline{C} C_L C_S \delta^2 \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)} \left\|\mathbb{D}(\boldsymbol{\eta})\right\|_{L^3(\Omega)} + \sum_{j=1}^{J} \beta \left\|\boldsymbol{\phi}^h \cdot \boldsymbol{\tau}_j\right\|_{L^2(\Gamma_j)} \left\|\boldsymbol{\eta} \cdot \boldsymbol{\tau}_j\right\|_{L^2(\Gamma_j)}
\end{aligned}
$$

with

$$
C_L = \max\left\{\left\|\mathbb{D}(\boldsymbol{w})\right\|_{L^3(\Omega)}, \left\|\mathbb{D}\left(\tilde{\boldsymbol{w}}^h\right)\right\|_{L^3(\Omega)}\right\}. \tag{8.133}
$$

The terms on the right-hand side are estimated further by Young's inequality (A.5) and the definition of the norm in V', (8.113),

$$
\begin{aligned}
& \overline{C} C_L C_S \delta^2 \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)} \left\|\mathbb{D}(\boldsymbol{\eta})\right\|_{L^3(\Omega)} \\
& \leq \frac{\underline{C} C_S \delta^2}{3} \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)}^3 + \frac{2\overline{C}^{3/2} C_L^{3/2} C_S \delta^2}{3\underline{C}^{1/2}} \left\|\mathbb{D}(\boldsymbol{\eta})\right\|_{L^3(\Omega)}^{3/2},
\end{aligned}
$$

$$
\mu \left\|\nabla \cdot \boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \left\|\nabla \cdot \boldsymbol{\eta}\right\|_{L^2(\Omega)} \leq \frac{\mu}{4} \left\|\nabla \cdot \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 + \mu \left\|\nabla \cdot \boldsymbol{\eta}\right\|_{L^2(\Omega)}^2,
$$

$$
\begin{aligned}
\left(\partial_t \boldsymbol{\eta}, \boldsymbol{\phi}^h\right) & \leq \left\|\partial_t \boldsymbol{\eta}\right\|_{V'} \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)} \\
& \leq \frac{2}{3} \left(\underline{C} C_S \delta^2\right)^{-1/2} \left\|\partial_t \boldsymbol{\eta}\right\|_{V'}^{3/2} + \frac{\underline{C} C_S \delta^2}{3} \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)}^3.
\end{aligned}
$$

Young's inequality is applied to the other terms with $p = q = 2$ and $t = 1$.

Inserting these estimates in (8.132), using the Cauchy–Schwarz inequality (A.10), and collecting terms gives in the case $\mu > 0$ estimate (8.130). ■

Lemma 8.114 (Estimate of the Nonlinear Convective Term) *It holds*

$$
\begin{aligned}
& \left| n_{\text{skew}}\left(\boldsymbol{w}, \boldsymbol{w}, \boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\boldsymbol{w}^h, \boldsymbol{w}^h, \boldsymbol{\phi}^h\right) \right| \\
& \leq \frac{1}{4} \left\|\nabla \boldsymbol{\eta}\right\|_{L^3(\Omega)}^2 + \frac{\varepsilon_1}{6} \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)}^3 + \frac{C}{3\varepsilon_1^{1/2}} \left\|\boldsymbol{w}\right\|_{L^2(\Omega)}^{3/2} \left\|\boldsymbol{\eta}\right\|_{L^6(\Omega)}^{3/2} \\
& \quad + \frac{1}{4} \left\|\boldsymbol{\eta}\right\|_{L^6(\Omega)}^2 + \frac{\varepsilon_1}{6} \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)}^3 + \frac{C}{3\varepsilon_1^{1/2}} \left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)}^{3/2} \left\|\boldsymbol{\eta}\right\|_{L^6(\Omega)}^{3/2} \\
& \quad + \frac{\nu_0(\delta)}{3} \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)}^2 + \frac{\mu}{8} \left\|\nabla \cdot \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2
\end{aligned}
$$

$$
+\left[\frac{1}{4}\|\boldsymbol{w}\|_{L^6(\Omega)}^2+\frac{1}{4}\left\|\nabla \boldsymbol{w}^h\right\|_{L^3(\Omega)}^2+\frac{C}{\nu_0(\delta)}\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^2\right.
$$
$$
\left.+\frac{C}{\nu_0(\delta)^{1/2}\mu^{3/2}}\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^3\right]\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2. \tag{8.134}
$$

Proof Again, the guidelines for deriving estimates are to get norms of $\boldsymbol{\phi}^h$ that can be absorbed from the left-hand side of (8.132) or to obtain $\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2$ for the application of Gronwall's lemma. The norms of $\boldsymbol{\eta}$ have to be bounded by the stability estimates given at the beginning of this section.

A straightforward calculation gives the decomposition, compare (6.64),

$$
\begin{aligned}
&n_{\text{skew}}\left(\boldsymbol{w},\boldsymbol{w},\boldsymbol{\phi}^h\right)-n_{\text{skew}}\left(\boldsymbol{w}^h,\boldsymbol{w}^h,\boldsymbol{\phi}^h\right) \\
&\quad = n_{\text{skew}}\left(\boldsymbol{w},\boldsymbol{\eta},\boldsymbol{\phi}^h\right)+n_{\text{skew}}\left(\boldsymbol{\eta},\boldsymbol{w}^h,\boldsymbol{\phi}^h\right)-n_{\text{skew}}\left(\boldsymbol{\phi}^h,\boldsymbol{w}^h,\boldsymbol{\phi}^h\right),
\end{aligned} \tag{8.135}
$$

where $n_{\text{skew}}\left(\boldsymbol{w},\boldsymbol{\phi}^h,\boldsymbol{\phi}^h\right)=0$ has been used, see (6.26).

The first term of (8.135) is bounded using Hölder's inequality (A.9), Korn's inequality (8.112), and Young's inequality (A.5)

$$
\begin{aligned}
&\left|n_{\text{skew}}\left(\boldsymbol{w},\boldsymbol{\eta},\boldsymbol{\phi}^h\right)\right| \\
&= \frac{1}{2}\left|n_{\text{conv}}n\left(\boldsymbol{w},\boldsymbol{\eta},\boldsymbol{\phi}^h\right)-n_{\text{conv}}\left(\boldsymbol{w},\boldsymbol{\phi}^h,\boldsymbol{\eta}\right)\right| \\
&\le \frac{1}{2}\left(\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}\|\nabla\boldsymbol{\eta}\|_{L^3(\Omega)}\|\boldsymbol{w}\|_{L^6(\Omega)}+\left\|\nabla\boldsymbol{\phi}^h\right\|_{L^3(\Omega)}\|\boldsymbol{w}\|_{L^2(\Omega)}\|\boldsymbol{\eta}\|_{L^6(\Omega)}\right) \\
&\le \frac{1}{2}\left(\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}\|\nabla\boldsymbol{\eta}\|_{L^3(\Omega)}\|\boldsymbol{w}\|_{L^6(\Omega)}+C\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)}\|\boldsymbol{w}\|_{L^2(\Omega)}\|\boldsymbol{\eta}\|_{L^6(\Omega)}\right) \\
&\le \frac{1}{4}\|\boldsymbol{w}\|_{L^6(\Omega)}^2\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2+\frac{1}{4}\|\nabla\boldsymbol{\eta}\|_{L^3(\Omega)}^2+\frac{\varepsilon_1}{6}\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)}^3 \\
&\quad +\frac{C}{3\varepsilon_1^{1/2}}\|\boldsymbol{w}\|_{L^2(\Omega)}^{3/2}\|\boldsymbol{\eta}\|_{L^6(\Omega)}^{3/2}.
\end{aligned} \tag{8.136}
$$

In the same way, one obtains

$$
\begin{aligned}
n_{\text{skew}}\left(\boldsymbol{\eta},\boldsymbol{w}^h,\boldsymbol{\phi}^h\right) &\le \frac{1}{4}\left\|\nabla\boldsymbol{w}^h\right\|_{L^3(\Omega)}^2\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2+\frac{1}{4}\|\boldsymbol{\eta}\|_{L^6(\Omega)}^2 \\
&\quad +\frac{\varepsilon_1}{6}\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)}^3+\frac{C}{3\varepsilon_1^{1/2}}\left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)}^{3/2}\|\boldsymbol{\eta}\|_{L^6(\Omega)}^{3/2}.
\end{aligned} \tag{8.137}
$$

The estimate of the third term starts with applying (6.22) and Hölder's inequality such that

$$\begin{aligned} n_{\text{skew}}\left(\boldsymbol{\phi}^h, \boldsymbol{w}^h, \boldsymbol{\phi}^h\right) &= \frac{1}{2} n_{\text{conv}}\left(\boldsymbol{\phi}^h, \boldsymbol{w}^h, \boldsymbol{\phi}^h\right) - \frac{1}{2} n_{\text{conv}}\left(\boldsymbol{\phi}^h, \boldsymbol{\phi}^h, \boldsymbol{w}^h\right) \\ &= n_{\text{conv}}\left(\boldsymbol{\phi}^h, \boldsymbol{w}^h, \boldsymbol{\phi}^h\right) + \frac{1}{2}\left(\nabla \cdot \boldsymbol{\phi}^h, \boldsymbol{\phi}^h \cdot \boldsymbol{w}^h\right) \\ &\le \left\|\nabla \boldsymbol{w}^h\right\|_{L^3(\Omega)} \left\|\boldsymbol{\phi}^h\right\|_{L^3(\Omega)}^2 + \frac{1}{2}\left|\left(\nabla \cdot \boldsymbol{\phi}^h, \boldsymbol{\phi}^h \cdot \boldsymbol{w}^h\right)\right| . \end{aligned} \tag{8.138}$$

In this estimate, the maximal regularity is used for $\nabla \boldsymbol{w}^h$ and both factors $\boldsymbol{\phi}^h$ are treated the same way. By the Sobolev imbedding $H^{1/2}(\Omega) \to L^3(\Omega)$, see (A.15), the interpolation theorem (A.13) and Poincare's inequality (A.12), one obtains

$$\begin{aligned} \left\|\boldsymbol{\phi}^h\right\|_{L^3(\Omega)}^2 &\le C \left\|\boldsymbol{\phi}^h\right\|_{H^{1/2}(\Omega)}^2 \le C \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \left\|\boldsymbol{\phi}^h\right\|_{H^1(\Omega)} \\ &\le C \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)} . \end{aligned}$$

Inserting this estimate in the previous estimate and applying Korn's and Young's inequalities gives

$$\begin{aligned} \left|n_{\text{skew}}\left(\boldsymbol{\phi}^h, \boldsymbol{w}^h, \boldsymbol{\phi}^h\right)\right| \le \frac{\nu_0(\delta)}{6} \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)}^2 &+ \frac{3C}{2\nu_0(\delta)} \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^2 \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 \\ &+ \frac{1}{2}\left|\left(\nabla \cdot \boldsymbol{\phi}^h, \boldsymbol{\phi}^h \cdot \boldsymbol{w}^h\right)\right| . \end{aligned}$$

The last term of the right-hand side of this inequality is estimated by Hölder's and by Young's inequality leading to

$$\begin{aligned} \left|\left(\nabla \cdot \boldsymbol{\phi}^h, \boldsymbol{\phi}^h \cdot \boldsymbol{w}^h\right)\right| &\le \left\|\nabla \cdot \boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \left\|\boldsymbol{\phi}^h\right\|_{L^{18/7}(\Omega)} \left\|\boldsymbol{w}^h\right\|_{L^9(\Omega)} \\ &\le \frac{\mu}{4} \left\|\nabla \cdot \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 + \frac{1}{\mu} \left\|\boldsymbol{\phi}^h\right\|_{L^{18/7}(\Omega)}^2 \left\|\boldsymbol{w}^h\right\|_{L^9(\Omega)}^2 . \end{aligned}$$

The Sobolev imbedding theorem $W^{1,3}(\Omega) \to L^9(\Omega)$, see (A.16), implies together with Poincare's and Korn's inequality

$$\left\|\boldsymbol{w}^h\right\|_{L^9(\Omega)}^2 \le C \left\|\boldsymbol{w}^h\right\|_{W^{1,3}(\Omega)}^2 \le C \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3}^2 .$$

The Sobolev imbedding theorem implies also $H^{1/3}(\Omega) \to L^{18/7}(\Omega)$, see (A.15). With the interpolation theorem (A.13), Poincare's and Korn's inequality, it follows that

$$\begin{aligned} \left\|\boldsymbol{\phi}^h\right\|_{L^{18/7}(\Omega)}^2 &\le C \left\|\boldsymbol{\phi}^h\right\|_{H^{1/3}(\Omega)}^2 \le C \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^{4/3} \left\|\boldsymbol{\phi}^h\right\|_{H^1(\Omega)}^{2/3} \\ &\le C \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^{4/3} \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)}^{2/3} . \end{aligned}$$

These bounds and Young's inequality yields

$$n_{\text{skew}}\left(\boldsymbol{\phi}^h, \boldsymbol{w}^h, \boldsymbol{\phi}^h\right) \leq \frac{\nu_0(\delta)}{6}\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)}^2 + \frac{C}{\nu_0(\delta)}\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^2\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2$$
$$+\frac{\mu}{8}\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 + \frac{\nu_0(\delta)}{6}\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)}^2$$
$$+\frac{C}{\nu_0(\delta)^{1/2}\mu^{3/2}}\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3}^3\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2.$$

Combining this estimate with (8.136) and (8.137) gives (8.134). ■

Remark 8.115 (On Inequalities (8.130) and (8.134)) The right-hand sides of these inequalities have to be estimated further. Because of the appearance of $\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2$ on the right-hand side of (8.134), Gronwall's lemma will be applied. In Lemmas 8.116 and 8.117 it will be shown that the terms on the right-hand side are sufficiently regular. The proofs use the stability bound for $\boldsymbol{w}$ and $\boldsymbol{w}^h$. □

Lemma 8.116 (Regularity in Time of the Term in the Brackets in (8.134)) *Assume $\mu > 0$ and $\nu_0(\delta) > 0$ for $\delta > 0$. Let*

$$\varkappa(t) := \frac{1}{4}\|\boldsymbol{w}\|_{L^6(\Omega)}^2 + \frac{1}{4}\left\|\nabla\boldsymbol{w}^h\right\|_{L^3(\Omega)}^2 + \frac{C}{\nu_0(\delta)}\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^2$$
$$+C\nu_0(\delta)^{-1/2}\mu^{-3/2}\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^3,$$

then there is a constant $C_1(\delta)$ independent of ν and h such that for $0 < T < \infty$

$$\|\varkappa(t)\|_{L^1(0,T)} \leq C_1(\delta).$$

Proof By the Sobolev imbedding $W^{1,3}(\Omega) \to L^6(\Omega)$, see (A.16), Poincaré's inequality (A.12), and Korn's inequality (8.112), one has $\|\boldsymbol{w}\|_{L^6(\Omega)} \leq C\|\mathbb{D}(\boldsymbol{w})\|_{L^3(\Omega)}$ which is bounded uniformly in ν by (8.121) and (8.122). By the stability estimates (8.125) and (8.126), it follows that $\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)} \in L^3(0,T)$ uniformly in ν and h. Since $L^3(0,T) \subset L^2(0,T)$, it is also $\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)} \in L^2(0,T)$, such that $\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^3 \in L^1(0,T)$ and $\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^2 \in L^1(0,T)$ uniformly in ν and h. ■

Lemma 8.117 (Regularity in Time of the Other Terms From the Right-Hand Side of (8.134)) *Under the assumptions of Lemma 8.116 there is a constant $C_2(\delta)$ independent of ν and h such that for $T \in (0,\infty)$*

$$\left(\left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)}^{3/2} + \|\boldsymbol{w}\|_{L^2(\Omega)}^{3/2}\right) \leq C_2(\delta).$$

Proof The statement of the lemma follows for $\|\boldsymbol{w}\|_{L^2(\Omega)}$ by the stability estimates (8.121) and (8.122) since

$$\int_0^T \|\boldsymbol{w}\|_{L^2(\Omega)}^{3/2}\, dt \le \|\boldsymbol{w}\|_{L^\infty(0,T;L^2(\Omega))}^{3/2}\, T \le C(\delta).$$

Similarly, this property follows for $\|\boldsymbol{w}^h\|_{L^2(\Omega)}$ from (8.125) and (8.126). ■

Theorem 8.118 (Uniform Finite Element Error Estimate for the Natural Regularity of the Solution and $\nu_0(\delta) > 0$**)** *Assume that* $\nabla \boldsymbol{w} \in L^3\left(0,T;L^3(\Omega)\right)$, $\partial_t \boldsymbol{w} \in L^{3/2}(0,T;V')$, $r \in L^2\left(0,T;L^2(\Omega)\right)$, $\mu > 0$, *and* $\nu_0(\delta) > 0$. *Then, the error* $\boldsymbol{w}-\boldsymbol{w}^h$ *satisfies for* $0 < T < \infty$

$$\begin{aligned}
&\left\|\boldsymbol{w}-\boldsymbol{w}^h\right\|_{L^\infty(0,T;L^2(\Omega))}^2 + \mu\left\|\nabla\cdot\left(\boldsymbol{w}-\boldsymbol{w}^h\right)\right\|_{L^2(0,T;L^2(\Omega))}^2 \\
&\quad + (\nu + C\nu_0(\delta))\left\|\mathbb{D}\left(\boldsymbol{w}-\boldsymbol{w}^h\right)\right\|_{L^2(0,T;L^2(\Omega))}^2 \\
&\quad + \delta^2\left\|\mathbb{D}\left(\boldsymbol{w}-\boldsymbol{w}^h\right)\right\|_{L^3(0,T;L^3(\Omega))}^3 + \sum_{j=1}^J \beta\left\|\left(\boldsymbol{w}-\boldsymbol{w}^h\right)\cdot\boldsymbol{\tau}_j\right\|_{L^2(0,T;L^2(\Gamma_j))}^2 \\
&\le C\exp\left(C_1(\delta)\right)\left\|\left(\boldsymbol{w}-\boldsymbol{w}^h\right)(0)\right\|_{L^2(\Omega)}^2 + C\inf_{\substack{\tilde{\boldsymbol{w}}^h\in V_{\mathrm{div}}^h\\ q^h\in Q^h}}\mathcal{F}\left(\boldsymbol{w}-\tilde{\boldsymbol{w}}^h, r-q^h,\delta\right)
\end{aligned}$$

with

$$\begin{aligned}
&\mathcal{F}\left(\boldsymbol{w}-\tilde{\boldsymbol{w}}^h, r-q^h,\delta\right) \\
&= \left\|\boldsymbol{w}-\tilde{\boldsymbol{w}}^h\right\|_{L^\infty(0,T;L^2(\Omega))}^2 + \delta^2\left\|\mathbb{D}\left(\boldsymbol{w}-\tilde{\boldsymbol{w}}^h\right)\right\|_{L^3(0,T;L^3(\Omega))}^3 \\
&\quad + \exp\left(C_1(\delta)\right)\Bigg[\left\|\left(\boldsymbol{w}-\tilde{\boldsymbol{w}}^h\right)(0)\right\|_{L^2(\Omega)}^2 + \delta^{-1}\left\|\partial_t\left(\boldsymbol{w}-\tilde{\boldsymbol{w}}^h\right)\right\|_{L^{3/2}(0,T;V')}^{3/2} \\
&\quad + \mu\left\|\nabla\cdot\left(\boldsymbol{w}-\tilde{\boldsymbol{w}}^h\right)\right\|_{L^2(0,T;L^2(\Omega))}^2 + (2\nu + \nu_0(\delta))\left\|\mathbb{D}\left(\boldsymbol{w}-\tilde{\boldsymbol{w}}^h\right)\right\|_{L^2(0,T;L^2(\Omega))}^2 \\
&\quad + C(\delta)\left\|\mathbb{D}\left(\boldsymbol{w}-\tilde{\boldsymbol{w}}^h\right)\right\|_{L^3(0,T;L^3(\Omega))}^{3/2} + \sum_{j=1}^J \beta\left\|\left(\boldsymbol{w}-\tilde{\boldsymbol{w}}^h\right)\cdot\boldsymbol{\tau}_j\right\|_{L^2(0,T;L^2(\Gamma_j))}^2 \\
&\quad + \left\|\nabla\left(\boldsymbol{w}-\tilde{\boldsymbol{w}}^h\right)\right\|_{L^2(0,T;L^3(\Omega))}^2 + C_2(\delta)\left\|\boldsymbol{w}-\tilde{\boldsymbol{w}}^h\right\|_{L^{3/2}(0,T;L^6(\Omega))}^{3/2} \\
&\quad + \left\|\boldsymbol{w}-\tilde{\boldsymbol{w}}^h\right\|_{L^2(0,T;L^6(\Omega))}^2 + \frac{1}{\mu}\left\|r-q^h\right\|_{L^2(0,T;L^2(\Omega))}^2\Bigg]
\end{aligned}$$

and $C_1(\delta)$ *and* $C_2(\delta)$ *defined in Lemmas 8.116 and 8.117.*

Proof The bound (8.134) is substituted into (8.130) yielding the differential inequality

$$
\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2+\frac{3\mu}{8}\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2+\left(\nu+\frac{\nu_0(\delta)}{6}\right)\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)}^2\\
&\quad+\left(\frac{\underline{C}C_S\delta^2}{3}-\frac{\varepsilon_1}{3}\right)\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)}^3+\sum_{j=1}^{J}\frac{\beta}{2}\left\|\boldsymbol{\phi}^h\cdot\boldsymbol{\tau}_j\right\|_{L^2(\Gamma_j)}^2\\
&\le\left[\frac{2}{3\left(\underline{C}C_S\right)^{1/2}\delta}\left\|\partial_t\boldsymbol{\eta}\right\|_{V'}^{3/2}+\mu\left\|\nabla\cdot\boldsymbol{\eta}\right\|_{L^2(\Omega)}^2+\frac{2\nu+\nu_0(\delta)}{2}\left\|\mathbb{D}(\boldsymbol{\eta})\right\|_{L^2(\Omega)}^2\right.\\
&\quad+\frac{2\overline{C}^{3/2}C_L^{3/2}C_S\delta^2}{3\underline{C}^{1/2}}\left\|\mathbb{D}(\boldsymbol{\eta})\right\|_{L^3(\Omega)}^{3/2}+\sum_{j=1}^{J}\frac{\beta}{2}\left\|\boldsymbol{\eta}\cdot\boldsymbol{\tau}_j\right\|_{L^2(\Gamma_j)}^2+\frac{1}{4}\left\|\nabla\boldsymbol{\eta}\right\|_{L^3(\Omega)}^2\\
&\quad+\frac{C}{3\varepsilon_1^{1/2}}\left(\left\|\boldsymbol{w}\right\|_{L^2(\Omega)}^{3/2}+\left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)}^{3/2}\right)\left\|\boldsymbol{\eta}\right\|_{L^6(\Omega)}^{3/2}+\frac{1}{4}\left\|\boldsymbol{\eta}\right\|_{L^6(\Omega)}^2 \qquad (8.139)\\
&\quad\left.+\frac{1}{\mu}\left\|r-q^h\right\|_{L^2(\Omega)}^2\right]+\left[\frac{1}{4}\left\|\boldsymbol{w}\right\|_{L^6(\Omega)}^2+\frac{1}{4}\left\|\nabla\boldsymbol{w}^h\right\|_{L^3(\Omega)}^2+\frac{C}{\nu_0(\delta)}\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^2\right.\\
&\quad\left.+\frac{C}{\nu_0(\delta)^{1/2}\mu^{3/2}}\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^3\right]\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2.
\end{aligned}
$$

Picking $\varepsilon_1=\underline{C}C_S\delta^2/2$ such that

$$
\frac{\underline{C}C_S\delta^2}{3}-\frac{\varepsilon_1}{3}=\frac{\underline{C}C_S\delta^2}{6}\quad\Longrightarrow\quad\frac{1}{\varepsilon_1^{1/2}}=\frac{C}{\delta},
$$

and multiplying with 2 yields

$$
\begin{aligned}
&\frac{d}{dt}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2+\frac{3\mu}{4}\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2+\left(2\nu+\frac{\nu_0(\delta)}{3}\right)\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)}^2\\
&\quad+\frac{\underline{C}C_S\delta^2}{3}\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)}^3+\sum_{j=1}^{J}\beta\left\|\boldsymbol{\phi}^h\cdot\boldsymbol{\tau}_j\right\|_{L^2(\Gamma_j)}^2\\
&\le\left[\frac{4}{3\left(\underline{C}C_S\right)^{1/2}\delta}\left\|\partial_t\boldsymbol{\eta}\right\|_{V'}^{3/2}+2\mu\left\|\nabla\cdot\boldsymbol{\eta}\right\|_{L^2(\Omega)}^2+\left(2\nu+\nu_0(\delta)\right)\left\|\mathbb{D}(\boldsymbol{\eta})\right\|_{L^2(\Omega)}^2\right.\\
&\quad+\frac{4\overline{C}^{3/2}C_L^{3/2}C_S\delta^2}{3\underline{C}^{1/2}}\left\|\mathbb{D}(\boldsymbol{\eta})\right\|_{L^3(\Omega)}^{3/2}+\sum_{j=1}^{J}\beta\left\|\boldsymbol{\eta}\cdot\boldsymbol{\tau}_j\right\|_{L^2(\Gamma_j)}^2+\frac{1}{2}\left\|\nabla\boldsymbol{\eta}\right\|_{L^3(\Omega)}^2
\end{aligned}
$$

$$
\begin{aligned}
&+C\delta^{-1}\left(\|\boldsymbol{w}\|_{L^2(\Omega)}^{3/2}+\left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)}^{3/2}\right)\|\boldsymbol{\eta}\|_{L^6(\Omega)}^{3/2}+\frac{1}{2}\|\boldsymbol{\eta}\|_{L^6(\Omega)}^2\\
&+\frac{2}{\mu}\left\|r-q^h\right\|_{L^2(\Omega)}^2\Bigg]+\Bigg[\frac{1}{2}\|\boldsymbol{w}\|_{L^6(\Omega)}^2+\frac{1}{2}\left\|\nabla\boldsymbol{w}^h\right\|_{L^3(\Omega)}^2\\
&+\frac{C}{\nu_0(\delta)}\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^2+C\nu_0(\delta)^{-1/2}\mu^{-3/2}\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^3\Bigg]\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2.
\end{aligned}
$$

Before applying Gronwall's lemma, Lemma A.55, one has to check that all functions are sufficiently smooth in time. All terms that involve only norms of $\boldsymbol{\eta}$ and derivatives of $\boldsymbol{\eta}$ are in $L^1(0,T)$. The other term in the first bracket is shown to be in $L^1(0,T)$ in Lemma 8.117. Finally, the term in the second bracket is also in $L^1(0,T)$ by Lemma 8.116. The application of Gronwall's lemma of form (A.40) gives for almost all $t\in[0,T]$

$$
\begin{aligned}
&\left\|\boldsymbol{\phi}^h(t)\right\|_{L^2(\Omega)}^2+\mu\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(0,t;L^2(\Omega))}^2+(\nu+C\nu_0(\delta))\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(0,t;L^2(\Omega))}^2\\
&\quad+\delta^2\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(0,t;L^3(\Omega))}^3+\sum_{j=1}^J\beta\left\|\boldsymbol{\phi}^h\cdot\boldsymbol{\tau}_j\right\|_{L^2\left(0,t;L^2\left(\Gamma_j\right)\right)}^2\\
&\le C\exp\left(\|\varkappa(t)\|_{L^1(0,t)}\right)\left\|\boldsymbol{\phi}^h(0)\right\|_{L^2(\Omega)}^2+C\exp\left(\|\varkappa(t)\|_{L^1(0,t)}\right)\\
&\quad\times\Bigg[\delta^{-1}\|\partial_t\boldsymbol{\eta}\|_{L^{3/2}(0,t;V')}^{3/2}+\mu\|\nabla\cdot\boldsymbol{\eta}\|_{L^2(0,t;L^2(\Omega))}^2+(2\nu+\nu_0(\delta))\|\mathbb{D}(\boldsymbol{\eta})\|_{L^2(0,t;L^2(\Omega))}^2\\
&\quad+\delta^2\int_0^t C_L^{3/2}\|\mathbb{D}(\boldsymbol{\eta})\|_{L^3(\Omega)}^{3/2}\,dt'+\sum_{j=1}^J\beta\left\|\boldsymbol{\eta}\cdot\boldsymbol{\tau}_j\right\|_{L^2\left(0,t;L^2\left(\Gamma_j\right)\right)}^2+\|\nabla\boldsymbol{\eta}\|_{L^2(0,t;L^3(\Omega))}^2\\
&\quad+\delta^{-1}\int_0^t\left(\|\boldsymbol{w}\|_{L^2(\Omega)}^{3/2}+\left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)}^{3/2}\right)\|\boldsymbol{\eta}\|_{L^6(\Omega)}^{3/2}\,dt'\\
&\quad+\|\boldsymbol{\eta}\|_{L^2(0,t;L^6(\Omega))}^2+\frac{1}{\mu}\left\|r-q^h\right\|_{L^2(0,t;L^2(\Omega))}^2\Bigg].
\end{aligned}
$$

Using the definition of C_L in (8.133), Lemma 8.100, the Cauchy–Schwarz inequality (A.10) in $L^2(0,T)$, and the stability estimates (8.121) and (8.122), one gets

$$
\begin{aligned}
&\int_0^t C_L^{3/2}\|\mathbb{D}(\boldsymbol{\eta})\|_{L^3(\Omega)}^{3/2}\,dt'\\
&\le\int_0^t\left(\max\left\{\|\mathbb{D}(\boldsymbol{w})\|_{L^3(\Omega)},\left\|\mathbb{D}\left(\tilde{\boldsymbol{w}}^h\right)\right\|_{L^3(\Omega)}\right\}\right)^{3/2}\|\mathbb{D}(\boldsymbol{\eta})\|_{L^3(\Omega)}^{3/2}\,dt'\\
&\le C\int_0^t\|\mathbb{D}(\boldsymbol{w})\|_{L^3(\Omega)}^{3/2}\|\mathbb{D}(\boldsymbol{\eta})\|_{L^3(\Omega)}^{3/2}\,dt'
\end{aligned}
$$

$$\leq C \left\| \mathbb{D}\left(\boldsymbol{w}\right) \right\|^{3/2}_{L^3\left(0,t;L^3(\Omega)\right)} \left\| \mathbb{D}\left(\boldsymbol{\eta}\right) \right\|^{3/2}_{L^3\left(0,t;L^3(\Omega)\right)}$$

$$\leq C\left(\delta\right) \left\| \mathbb{D}\left(\boldsymbol{\eta}\right) \right\|^{3/2}_{L^3\left(0,t;L^3(\Omega)\right)}.$$

All constants which does not depend on the problem or the discretization are collected in the generic constants C. Using the definition of $C_2(\delta)$, Lemma 8.117, and applying the essential supremum on $(0, T)$ on both sides of the inequality gives

$$\begin{aligned}
&\left\| \boldsymbol{\phi}^h \right\|^2_{L^\infty\left(0,T;L^2(\Omega)\right)} + \mu \left\| \nabla \cdot \boldsymbol{\phi}^h \right\|^2_{L^2\left(0,T;L^2(\Omega)\right)} \\
&\quad + \left(\nu + C\nu_0\left(\delta\right)\right) \left\| \mathbb{D}\left(\boldsymbol{\phi}^h\right) \right\|^2_{L^2\left(0,T;L^2(\Omega)\right)} + \delta^2 \left\| \mathbb{D}\left(\boldsymbol{\phi}^h\right) \right\|^3_{L^3\left(0,T;L^3(\Omega)\right)} \\
&\quad + \sum_{j=1}^{J} \beta \left\| \boldsymbol{\phi}^h \cdot \boldsymbol{\tau}_j \right\|^2_{L^2\left(0,T;L^2\left(\Gamma_j\right)\right)} \\
&\leq C \exp\left(\left\| \varkappa\left(t\right) \right\|_{L^1(0,T)}\right) \left\| \boldsymbol{\phi}^h\left(0\right) \right\|^2_{L^2(\Omega)} \\
&\quad + C \exp\left(\left\| \varkappa\left(t\right) \right\|_{L^1(0,T)}\right) \Bigg[\delta^{-1} \left\| \partial_t \boldsymbol{\eta} \right\|^{3/2}_{L^{3/2}(0,T;V')} + \mu \left\| \nabla \cdot \boldsymbol{\eta} \right\|^2_{L^2\left(0,T;L^2(\Omega)\right)} \\
&\quad + \left(2\nu + \nu_0\left(\delta\right)\right) \left\| \mathbb{D}\left(\boldsymbol{\eta}\right) \right\|^2_{L^2\left(0,T;L^2(\Omega)\right)} + C\left(\delta\right) \left\| \mathbb{D}\left(\boldsymbol{\eta}\right) \right\|^{3/2}_{L^3\left(0,T;L^3(\Omega)\right)} \\
&\quad + \sum_{j=1}^{J} \beta \left\| \boldsymbol{\eta} \cdot \boldsymbol{\tau}_j \right\|^2_{L^2\left(0,T;L^2\left(\Gamma_j\right)\right)} + \left\| \nabla \boldsymbol{\eta} \right\|^2_{L^2\left(0,T;L^3(\Omega)\right)} \\
&\quad + C_2\left(\delta\right) \left\| \boldsymbol{\eta} \right\|^{3/2}_{L^{3/2}\left(0,T;L^6(\Omega)\right)} + \left\| \boldsymbol{\eta} \right\|^2_{L^2\left(0,T;L^6(\Omega)\right)} + \frac{1}{\mu} \left\| r - q^h \right\|^2_{L^2\left(0,T;L^2(\Omega)\right)} \Bigg].
\end{aligned}$$

The triangle inequality implies

$$\begin{aligned}
&\left\| \boldsymbol{w} - \boldsymbol{w}^h \right\|^2_{L^\infty\left(0,T;L^2(\Omega)\right)} + \mu \left\| \nabla \cdot \left(\boldsymbol{w} - \boldsymbol{w}^h\right) \right\|^2_{L^2\left(0,T;L^2(\Omega)\right)} \\
&\quad + \left(\nu + C\nu_0\left(\delta\right)\right) \left\| \mathbb{D}\left(\boldsymbol{w} - \boldsymbol{w}^h\right) \right\|^2_{L^2\left(0,T;L^2(\Omega)\right)} \\
&\quad + \delta^2 \left\| \mathbb{D}\left(\boldsymbol{w} - \boldsymbol{w}^h\right) \right\|^3_{L^3\left(0,T;L^3(\Omega)\right)} + \sum_{j=1}^{J} \beta \left\| \left(\boldsymbol{w} - \boldsymbol{w}^h\right) \cdot \boldsymbol{\tau}_j \right\|^2_{L^2\left(0,T;L^2\left(\Gamma_j\right)\right)} \\
&\leq C \Bigg[\left\| \boldsymbol{\eta} \right\|^2_{L^\infty\left(0,T;L^2(\Omega)\right)} + \mu \left\| \nabla \cdot \boldsymbol{\eta} \right\|^2_{L^2\left(0,T;L^2(\Omega)\right)} \\
&\quad + \left(\nu + C\nu_0\left(\delta\right)\right) \left\| \mathbb{D}\left(\boldsymbol{\eta}\right) \right\|^2_{L^2\left(0,T;L^2(\Omega)\right)} + \delta^2 \left\| \mathbb{D}\left(\boldsymbol{\eta}\right) \right\|^3_{L^3\left(0,T;L^3(\Omega)\right)}
\end{aligned}$$

$$
\begin{aligned}
&+\sum_{j=1}^{J}\beta\left\|\boldsymbol{\eta}\cdot\boldsymbol{\tau}_j\right\|^2_{L^2\left(0,T;L^2\left(\Gamma_j\right)\right)}+\left\|\boldsymbol{\phi}^h\right\|^2_{L^\infty\left(0,T;L^2(\Omega)\right)}\\
&+\mu\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|^2_{L^2\left(0,T;L^2(\Omega)\right)}+\left(\nu+C\nu_0\left(\delta\right)\right)\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|^2_{L^2\left(0,T;L^2(\Omega)\right)}\\
&\left.+\delta^2\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|^3_{L^3\left(0,T;L^3(\Omega)\right)}+\sum_{j=1}^{J}\beta\left\|\boldsymbol{\phi}^h\cdot\boldsymbol{\tau}_j\right\|^2_{L^2\left(0,T;L^2\left(\Gamma_j\right)\right)}\right].
\end{aligned}
$$

With the estimate for $\boldsymbol{\phi}^h$ and since $\exp\left(\left\|\varkappa\left(t\right)\right\|_{L^1(0,T)}\right)\geq 1$, one gets

$$
\begin{aligned}
&\left\|\boldsymbol{w}-\boldsymbol{w}^h\right\|^2_{L^\infty\left(0,T;L^2(\Omega)\right)}+\mu\left\|\nabla\cdot\left(\boldsymbol{w}-\boldsymbol{w}^h\right)\right\|^2_{L^2\left(0,T;L^2(\Omega)\right)}\\
&\quad+\left(\nu+C\nu_0\left(\delta\right)\right)\left\|\mathbb{D}\left(\boldsymbol{w}-\boldsymbol{w}^h\right)\right\|^2_{L^2\left(0,T;L^2(\Omega)\right)}\\
&\quad+\delta^2\left\|\mathbb{D}\left(\boldsymbol{w}-\boldsymbol{w}^h\right)\right\|^3_{L^3\left(0,T;L^3(\Omega)\right)}+\sum_{j=1}^{J}\beta\left\|\left(\boldsymbol{w}-\boldsymbol{w}^h\right)\cdot\boldsymbol{\tau}_j\right\|^2_{L^2\left(0,T;L^2\left(\Gamma_j\right)\right)}\\
&\leq C\exp\left(\left\|\varkappa\left(t\right)\right\|_{L^1(0,T)}\right)\left\|\boldsymbol{\phi}^h\left(0\right)\right\|^2_{L^2(\Omega)}\\
&\quad+C\left\{\left\|\boldsymbol{\eta}\right\|^2_{L^\infty\left(0,T;L^2(\Omega)\right)}+\delta^2\left\|\mathbb{D}\left(\boldsymbol{\eta}\right)\right\|^3_{L^3\left(0,T;L^3(\Omega)\right)}\right.\\
&\quad+\exp\left(\left\|\varkappa\left(t\right)\right\|_{L^1(0,T)}\right)\left[\delta^{-1}\left\|\partial_t\boldsymbol{\eta}\right\|^{3/2}_{L^{3/2}(0,T;V')}+\mu\left\|\nabla\cdot\boldsymbol{\eta}\right\|^2_{L^2\left(0,T;L^2(\Omega)\right)}\right.\\
&\quad+\left(2\nu+\nu_0\left(\delta\right)\right)\left\|\mathbb{D}\left(\boldsymbol{\eta}\right)\right\|^2_{L^2\left(0,T;L^2(\Omega)\right)}+C\left(\delta\right)\left\|\mathbb{D}\left(\boldsymbol{\eta}\right)\right\|^{3/2}_{L^3\left(0,T;L^3(\Omega)\right)}\\
&\quad+\sum_{j=1}^{J}\beta\left\|\boldsymbol{\eta}\cdot\boldsymbol{\tau}_j\right\|^2_{L^2\left(0,T;L^2\left(\Gamma_j\right)\right)}+\left\|\nabla\boldsymbol{\eta}\right\|^2_{L^2\left(0,T;L^3(\Omega)\right)}\\
&\quad\left.\left.+C_2\left(\delta\right)\left\|\boldsymbol{\eta}\right\|^{3/2}_{L^{3/2}\left(0,T;L^6(\Omega)\right)}+\left\|\boldsymbol{\eta}\right\|^2_{L^2\left(0,T;L^6(\Omega)\right)}+\frac{1}{\mu}\left\|r-q^h\right\|^2_{L^2\left(0,T;L^2(\Omega)\right)}\right]\right\}.
\end{aligned}
$$

Applying the triangle inequality

$$
\left\|\boldsymbol{\phi}^h\left(0\right)\right\|^2_{L^2(\Omega)}\leq C\left(\left\|\left(\boldsymbol{w}-\boldsymbol{w}^h\right)\left(0\right)\right\|^2_{L^2(\Omega)}+\left\|\boldsymbol{\eta}\left(0\right)\right\|^2_{L^2(\Omega)}\right)
$$

and taking the infimum over $\tilde{\boldsymbol{w}}^h$ completes the proof of Theorem 8.118. ■

Remark 8.119 (The Case $\nabla\boldsymbol{w}\in L^3\left(0,T;L^3\left(\Omega\right)\right)$ and $\nu_0\left(\delta\right)=0$) The result proved in Theorem 8.118 requires that $\nu_0\left(\delta\right)>0$, i.e., there is some artificial viscosity in addition to the Smagorinsky model. In practice, the Smagorinsky model is used without artificial viscosity such that an error analysis for the case $\nu_0(\delta) = 0$ is of much interest. However, so far uniform finite element error estimates for the case $\nabla \boldsymbol{w} \in L^3\left(0, T; L^3(\Omega)\right)$ and $\nu_0(\delta) = 0$ are not known. This remark explains why a straightforward approach for deriving an error estimate fails.

Let $\nu_0(\delta) = 0$ and $\nabla \boldsymbol{w} \in L^3\left(0, T; L^3(\Omega)\right)$. The key of the error analysis is the estimate of the nonlinear convective term. Using the decomposition (8.135), the first and the second term are estimated as before, see (8.136) and (8.137). The critical term is the last one. As in estimate (8.138), the maximal regularity will be used for $\nabla \boldsymbol{w}^h$. However, the two factors $\boldsymbol{\phi}^h$ will be treated in a different way, having mind that $\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2$ might be a good term on the right-hand side for the application of Gronwall's lemma. Bounding this term by Hölder's inequality (A.9), the Sobolev imbedding $W^{1,3}(\Omega) \to L^6(\Omega)$, see (A.16), Korn's inequality (8.112), and Young's inequality (A.5) yields

$$
\begin{aligned}
& n_{\text{skew}}\left(\boldsymbol{\phi}^h, \boldsymbol{w}^h, \boldsymbol{\phi}^h\right) \\
& = \frac{1}{2}\left(n_{\text{conv}}\left(\boldsymbol{\phi}^h, \boldsymbol{w}^h, \boldsymbol{\phi}^h\right) - n_{\text{conv}}\left(\boldsymbol{\phi}^h, \boldsymbol{\phi}^h, \boldsymbol{w}^h\right)\right) \\
& \le \frac{1}{2}\left(\left\|\nabla \boldsymbol{w}^h\right\|_{L^3(\Omega)} \left\|\boldsymbol{\phi}^h\right\|_{L^6(\Omega)} \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} + \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^3(\Omega)} \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \left\|\boldsymbol{w}^h\right\|_{L^6(\Omega)} \right) \\
& \le C \left\|\nabla \boldsymbol{w}^h\right\|_{L^3(\Omega)} \left\|\boldsymbol{\phi}^h\right\|_{W^{1,3}(\Omega)} \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} + C \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)} \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \left\|\boldsymbol{w}^h\right\|_{W^{1,3}(\Omega)} \\
& \le \frac{\varepsilon_1}{6} \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)}^3 + \frac{C}{\varepsilon_1^{1/2}} \left\|\nabla \boldsymbol{w}^h\right\|_{L^3(\Omega)}^{3/2} \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^{3/2} \\
& \quad + \frac{\varepsilon_1}{6} \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)}^3 + \frac{C}{\varepsilon_1^{1/2}} \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^{3/2} \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^{3/2},
\end{aligned}
\tag{8.140}
$$

where in the last step Poincaré's inequality (A.12) and Korn's inequality were used to obtain

$$
\left\|\boldsymbol{\phi}^h\right\|_{W^{1,3}(\Omega)} \le C \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^3(\Omega)} \le C \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)}.
$$

Collecting the terms of (8.136), (8.137), and (8.140) yields

$$
\begin{aligned}
& \left| n_{\text{skew}}\left(\boldsymbol{w}, \boldsymbol{w}, \boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\boldsymbol{w}^h, \boldsymbol{w}^h, \boldsymbol{\phi}^h\right) \right| \\
& \le \left[\frac{1}{4} \left\|\nabla \boldsymbol{\eta}\right\|_{L^3(\Omega)}^2 + \frac{C}{\varepsilon_1^{1/2}} \left(\left\|\boldsymbol{w}\right\|_{L^2(\Omega)}^{3/2} + \left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)}^{3/2} \right) \left\|\boldsymbol{\eta}\right\|_{L^6(\Omega)}^{3/2} + \frac{1}{4} \left\|\boldsymbol{\eta}\right\|_{L^6(\Omega)}^2 \right] \\
& \quad + \frac{2\varepsilon_1}{3} \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)}^3 + \frac{C}{\varepsilon_1^{1/2}} \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^{3/2} \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^{3/2}
\end{aligned}
\tag{8.141}
$$

$$+\left[\frac{1}{4}\|\boldsymbol{w}\|_{L^6}^2+\frac{1}{4}\left\|\nabla \boldsymbol{w}^h\right\|_{L^3}^2\right]\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 .$$

Choosing now

$$\frac{2\varepsilon_1}{3}=\frac{1}{6}\underline{C}C_S\delta^2 \quad\Longrightarrow\quad \varepsilon_1=\mathcal{O}\left(\delta^2\right),\ \frac{1}{\varepsilon_1^{1/2}}=\mathcal{O}\left(\delta^{-1}\right),$$

one obtains with (8.130) the differential inequality

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2+\frac{\mu}{2}\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2+\nu\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)}^2\\
&\quad+\frac{\underline{C}C_S\delta^2}{6}\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)}^3+\sum_{j=1}^{J}\frac{\beta}{2}\left\|\boldsymbol{\phi}^h\cdot\boldsymbol{\tau}_j\right\|_{L^2(\Gamma_j)}^2\\
&\le\left[\frac{2}{3\left(\underline{C}C_S\right)^{1/2}\delta}\left\|\partial_t\boldsymbol{\eta}\right\|_{V'}^{3/2}+\mu\left\|\nabla\cdot\boldsymbol{\eta}\right\|_{L^2(\Omega)}^2+\nu\left\|\mathbb{D}\left(\boldsymbol{\eta}\right)\right\|_{L^2(\Omega)}^2\right.\\
&\quad+\frac{2\overline{C}^{3/2}C_L^{3/2}C_S\delta^2}{3\underline{C}^{1/2}}\left\|\mathbb{D}\left(\boldsymbol{\eta}\right)\right\|_{L^3(\Omega)}^{3/2}+\sum_{j=1}^{J}\frac{\beta}{2}\left\|\boldsymbol{\eta}\cdot\boldsymbol{\tau}_j\right\|_{L^2(\Gamma_j)}^2+\frac{1}{4}\left\|\nabla\boldsymbol{\eta}\right\|_{L^3(\Omega)}^2\\
&\quad+C\delta^{-1}\left(\left\|\boldsymbol{w}\right\|_{L^2(\Omega)}^{3/2}+\left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)}^{3/2}\right)\left\|\boldsymbol{\eta}\right\|_{L^6(\Omega)}^{3/2}+\frac{1}{4}\left\|\boldsymbol{\eta}\right\|_{L^6(\Omega)}^2\\
&\quad\left.+\frac{1}{\mu}\left\|r-q^h\right\|_{L^2(\Omega)}^2\right]+C\delta^{-1}\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^{3/2}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^{3/2}\\
&\quad+\left[\frac{1}{4}\|\boldsymbol{w}\|_{L^6}^2+\frac{1}{4}\left\|\nabla\boldsymbol{w}^h\right\|_{L^3}^2\right]\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 .
\end{aligned}\tag{8.142}$$

The last step of the proof would be the application of Gronwall's lemma, Lemma A.55, for $f(t) = \left\|\boldsymbol{\phi}^h(t)\right\|_{L^2(\Omega)}^2$. However, the term with $\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^{3/2} = \left(\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2\right)^{3/4}$ in the right-hand side of (8.142) does not fit into the basic inequality (A.38) of Lemma A.55. The power $3/4$ is too small. Thus, Gronwall's lemma cannot be applied and this approach fails.

- Also other attempts for deriving a uniform error estimate in the case $\nabla\boldsymbol{w} \in L^3\left(0,T;L^3(\Omega)\right)$ and $\nu_0(\delta) = 0$ failed so far, see John and Layton (2002) and John (2004, Sect. 8.1.4).
- It would be possible to prove a finite element error estimate that is not uniform, i.e., where the constant depends on ν, see John (2004, Sect. 8.1.7) for a discussion of this topic.

- A uniform error bound in the case $\nu_0(\delta) \geq 0$ can be achieved if a higher regularity of the solution is assumed, see Theorem 8.120. This situation is similar to finite element error analysis of the Navier–Stokes equations, where higher regularity assumptions also enable the derivation of error bounds with a weaker dependency on the viscosity, compare Remark 7.39.

Maybe, the deeper reason for the failure is that one cannot exclude that the Smagorinsky model with $\nu_0(\delta) = 0$ vanishes in some points $(t, \boldsymbol{x})$ or even in some regions. Thus, from the analytical point of view, there is no uniform positive bound form below for the additional viscosity and maybe one cannot expect to obtain different estimates than for the Navier–Stokes equations since the analysis uses only global estimates. □

Theorem 8.120 (Uniform Finite Element Error Estimate for the Case $\nu_0(\delta) \geq 0$ and Higher Regularity of $\boldsymbol{w}$ Uniformly in ν) *Suppose $\nu_0(\delta) \geq 0, \mu > 0$, and $\boldsymbol{w} \in L^2\left(0, T; W^{1,\infty}(\Omega)\right)$, $\nabla \boldsymbol{w} \in L^4\left(0, T; L^3(\Omega)\right)$, both uniformly in ν. Let*

$$\varkappa(t) := \frac{3}{4} + \|\nabla \boldsymbol{w}\|_{L^\infty(\Omega)} + \left(\frac{1}{4} + \frac{1}{4\mu}\right) \|\boldsymbol{w}\|_{L^\infty(\Omega)}^2 + \frac{1}{2} \|\nabla \boldsymbol{w}\|_{L^\infty(\Omega)}^2,$$

then there is a $C_3 = C_3(\boldsymbol{w})$ such that $\|\varkappa(t)\|_{L^1(0,T)} \leq C_3(\boldsymbol{w})$. Let $C_4 = C_4(\delta)$ be such that $\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3\left(0,T;L^3\right)} \leq C_4(\delta)$. Then, the error $\boldsymbol{w} - \boldsymbol{w}^h$ satisfies

$$\begin{aligned}
&\left\|\boldsymbol{w} - \boldsymbol{w}^h\right\|_{L^\infty\left(0,T;L^2(\Omega)\right)}^2 + \delta^2 \left\|\mathbb{D}\left(\boldsymbol{w} - \boldsymbol{w}^h\right)\right\|_{L^3\left(0,T;L^3(\Omega)\right)}^3 \\
&\quad + \mu \left\|\nabla \cdot \left(\boldsymbol{w} - \boldsymbol{w}^h\right)\right\|_{L^2\left(0,T;L^2(\Omega)\right)}^2 + (\nu + C\nu_0(\delta)) \left\|\mathbb{D}\left(\boldsymbol{w} - \boldsymbol{w}^h\right)\right\|_{L^2\left(0,T;L^2(\Omega)\right)}^2 \\
&\quad + \sum_{j=1}^{J} \beta \left\|\left(\boldsymbol{w} - \boldsymbol{w}^h\right) \cdot \hat{\boldsymbol{\tau}}_j\right\|_{L^2\left(0,T;L^2(\Gamma_j)\right)}^2 \qquad (8.143) \\
&\leq C \exp\left(C_3(\boldsymbol{w})\right) \left\|\left(\boldsymbol{w} - \boldsymbol{w}^h\right)(0)\right\|_{L^2(\Omega)}^2 + C \inf_{\substack{\tilde{\boldsymbol{w}}^h \in V_{\mathrm{div}}^h \\ q^h \in Q^h}} \mathcal{F}\left(\boldsymbol{w} - \tilde{\boldsymbol{w}}^h, r - q^h, \delta\right)
\end{aligned}$$

with

$$\begin{aligned}
&\mathcal{F}\left(\boldsymbol{w} - \tilde{\boldsymbol{w}}^h, r - q^h, \delta\right) \qquad (8.144) \\
&= \left\|\boldsymbol{w} - \tilde{\boldsymbol{w}}^h\right\|_{L^\infty\left(0,T;L^2(\Omega)\right)}^2 + \delta^2 \left\|\mathbb{D}\left(\boldsymbol{w} - \tilde{\boldsymbol{w}}^h\right)\right\|_{L^3\left(0,T;L^3(\Omega)\right)}^3 \\
&\quad + \exp\left(C_3(\boldsymbol{w})\right) \Bigg[\left\|\left(\boldsymbol{w} - \tilde{\boldsymbol{w}}^h\right)(0)\right\|_{L^2(\Omega)}^2 + \delta^{-1} \left\|\partial_t \left(\boldsymbol{w} - \tilde{\boldsymbol{w}}^h\right)\right\|_{L^{3/2}(0,T;V')}^{3/2} \\
&\quad + (2\nu + \nu_0(\delta)) \left\|\mathbb{D}\left(\boldsymbol{w} - \tilde{\boldsymbol{w}}^h\right)\right\|_{L^2\left(0,T;L^2(\Omega)\right)}^2
\end{aligned}$$

$$+\sum_{j=1}^{J}\beta\left\|\left(\boldsymbol{w}-\tilde{\boldsymbol{w}}^h\right)\cdot\hat{\boldsymbol{\tau}}_j\right\|^2_{L^2\left(0,T;L^2\left(\Gamma_j\right)\right)}+C\left(\delta\right)\left\|\mathbb{D}\left(\boldsymbol{w}-\tilde{\boldsymbol{w}}^h\right)\right\|^{3/2}_{L^3\left(0,T;L^3(\Omega)\right)}$$

$$+\mu^{-1}\left\|r-q^h\right\|^2_{L^2\left(0,T;L^2(\Omega)\right)}+\left(\frac{1}{4}+\mu\right)\left\|\nabla\cdot\left(\boldsymbol{w}-\tilde{\boldsymbol{w}}^h\right)\right\|^2_{L^2\left(0,T;L^2(\Omega)\right)}$$

$$+\left\|\boldsymbol{w}-\tilde{\boldsymbol{w}}^h\right\|^2_{L^2\left(0,T,L^2(\Omega)\right)}$$

$$\left.+C_4\left(\delta\right)\left(\left\|\mathbb{D}\left(\boldsymbol{w}-\tilde{\boldsymbol{w}}^h\right)\right\|^2_{L^{18/5}\left(0,T;L^3(\Omega)\right)}+\left\|\boldsymbol{w}-\tilde{\boldsymbol{w}}^h\right\|^2_{L^6\left(0,T;L^6(\Omega)\right)}\right)\right].$$

Proof The proof starts with using a special case of the Gagliardo–Nirenberg inequality, e.g., see Nirenberg (1959),

$$\|\boldsymbol{w}\|_{L^6(\Omega)}\le C\,\|\nabla\boldsymbol{w}\|^{2/3}_{L^3(\Omega)}\,\|\boldsymbol{w}\|^{1/3}_{L^2(\Omega)}\,, \tag{8.145}$$

which holds because $\boldsymbol{w}$ vanishes on Γ_0, see also the discussion of this inequality in John and Layton (2002). Applying (8.145) and Korn's inequality (3.43), it follows that

$$\begin{aligned}\|\boldsymbol{w}\|^6_{L^6\left(0,T;L^6(\Omega)\right)}&=\int_0^T\|\boldsymbol{w}\|^6_{L^6(\Omega)}\,dt\le C\int_0^T\|\nabla\boldsymbol{w}\|^4_{L^3(\Omega)}\,\|\boldsymbol{w}\|^2_{L^2(\Omega)}\,dt\\&\le C\,\|\boldsymbol{w}\|^2_{L^\infty\left(0,T;L^2(\Omega)\right)}\int_0^T\|\mathbb{D}\left(\boldsymbol{w}\right)\|^4_{L^3(\Omega)}\,dt\\&=C\,\|\boldsymbol{w}\|^2_{L^\infty\left(0,T;L^2(\Omega)\right)}\,\|\mathbb{D}\left(\boldsymbol{w}\right)\|^4_{L^4\left(0,T;L^3(\Omega)\right)}\le C<\infty\end{aligned}$$

by the regularity assumptions and stability estimates (8.121) and (8.122). Note also that $\nabla\boldsymbol{w}\in L^4\left(0,T;L^3\left(\Omega\right)\right)$ uniformly in ν implies $\nabla\boldsymbol{w}\in L^{18/5}\left(0,T;L^3\left(\Omega\right)\right)$ uniformly in ν since the time interval is bounded.

The proof is based on the differential inequality (8.130). The nonlinear convective term is decomposed in the form (6.65)

$$\begin{aligned}&\left|n_{\text{skew}}\left(\boldsymbol{w},\boldsymbol{w},\boldsymbol{\phi}^h\right)-n_{\text{skew}}\left(\boldsymbol{w}^h,\boldsymbol{w}^h,\boldsymbol{\phi}^h\right)\right|\\&=\left|n_{\text{skew}}\left(\boldsymbol{\eta}-\boldsymbol{\phi}^h,\boldsymbol{w},\boldsymbol{\phi}^h\right)+n_{\text{skew}}\left(\boldsymbol{w}^h,\boldsymbol{\eta}-\boldsymbol{\phi}^h,\boldsymbol{\phi}^h\right)\right|\\&=\left|n_{\text{skew}}\left(\boldsymbol{\eta},\boldsymbol{w},\boldsymbol{\phi}^h\right)-n_{\text{skew}}\left(\boldsymbol{\phi}^h,\boldsymbol{w},\boldsymbol{\phi}^h\right)+n_{\text{skew}}\left(\boldsymbol{w}^h,\boldsymbol{\eta},\boldsymbol{\phi}^h\right)\right|,\end{aligned}$$

where $n_{\text{skew}}\left(\boldsymbol{w}^h,\boldsymbol{\phi}^h,\boldsymbol{\phi}^h\right)=0$, see (6.26), was used. With this decomposition, the critical term is $n_{\text{skew}}\left(\boldsymbol{\phi}^h,\boldsymbol{w},\boldsymbol{\phi}^h\right)$ and not $n_{\text{skew}}\left(\boldsymbol{\phi}^h,\boldsymbol{w}^h,\boldsymbol{\phi}^h\right)$ as in (8.135). For estimating the critical term, the assumptions on $\boldsymbol{w}$ can be used now. Note that one can assume regularity for the solution of the continuous problem but not for the finite element solution.

The individual terms of the right-hand side are first transformed with (6.22) and then bounded by Hölder's inequality (A.9) and by Young's inequality (A.5)

$$\begin{aligned}&\left|n_{\text{skew}}\left(\boldsymbol{\eta},\boldsymbol{w},\boldsymbol{\phi}^h\right)\right| = \left|n_{\text{conv}}\left(\boldsymbol{\eta},\boldsymbol{w},\boldsymbol{\phi}^h\right) + \frac{1}{2}\left(\nabla\cdot\boldsymbol{\eta},\boldsymbol{\phi}^h\cdot\boldsymbol{w}\right)\right|\\
&\le \|\nabla\boldsymbol{w}\|_{L^\infty(\Omega)}\|\boldsymbol{\eta}\|_{L^2(\Omega)}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} + \frac{1}{2}\|\boldsymbol{w}\|_{L^\infty(\Omega)}\|\nabla\cdot\boldsymbol{\eta}\|_{L^2(\Omega)}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}\\
&\le \frac{1}{2}\|\boldsymbol{\eta}\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\nabla\boldsymbol{w}\|_{L^\infty(\Omega)}^2\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 + \frac{1}{4}\|\nabla\cdot\boldsymbol{\eta}\|_{L^2(\Omega)}^2\\
&\quad + \frac{1}{4}\|\boldsymbol{w}\|_{L^\infty(\Omega)}^2\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2,\end{aligned}$$

$$\begin{aligned}&\left|n_{\text{skew}}\left(\boldsymbol{\phi}^h,\boldsymbol{w},\boldsymbol{\phi}^h\right)\right| = \left|n_{\text{conv}}\left(\boldsymbol{\phi}^h,\boldsymbol{w},\boldsymbol{\phi}^h\right) + \frac{1}{2}\left(\nabla\cdot\boldsymbol{\phi}^h,\boldsymbol{w}\cdot\boldsymbol{\phi}^h\right)\right|\\
&\le \|\nabla\boldsymbol{w}\|_{L^\infty(\Omega)}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 + \frac{1}{2}\|\boldsymbol{w}\|_{L^\infty(\Omega)}\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}\\
&\le \|\nabla\boldsymbol{w}\|_{L^\infty(\Omega)}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 + \frac{\mu}{4}\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 + \frac{1}{4\mu}\|\boldsymbol{w}\|_{L^\infty(\Omega)}^2\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2,\end{aligned}$$

$$\begin{aligned}&\left|n_{\text{skew}}\left(\boldsymbol{w}^h,\boldsymbol{\eta},\boldsymbol{\phi}^h\right)\right| = \left|n_{\text{conv}}\left(\boldsymbol{w}^h,\boldsymbol{\eta},\boldsymbol{\phi}^h\right) + \frac{1}{2}\left(\nabla\cdot\boldsymbol{w}^h,\boldsymbol{\eta}\cdot\boldsymbol{\phi}^h\right)\right|\\
&\le \left\|\boldsymbol{w}^h\right\|_{L^6(\Omega)}\|\nabla\boldsymbol{\eta}\|_{L^3(\Omega)}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} + \frac{1}{2}\left\|\nabla\cdot\boldsymbol{w}^h\right\|_{L^3(\Omega)}\|\boldsymbol{\eta}\|_{L^6(\Omega)}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}\\
&\le \frac{1}{2}\left\|\boldsymbol{w}^h\right\|_{L^6(\Omega)}^2\|\nabla\boldsymbol{\eta}\|_{L^3(\Omega)}^2 + C\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^2\|\boldsymbol{\eta}\|_{L^6(\Omega)}^2 + \frac{3}{4}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2,\end{aligned}$$

where Korn's inequality (8.112) was used in the last line. The term $\left\|\boldsymbol{w}^h\right\|_{L^6(\Omega)}$ is bounded using the Gagliardo–Nirenberg inequality (8.145), Korn's inequality, and the uniform boundedness of $\left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)}$

$$\left\|\boldsymbol{w}^h\right\|_{L^6(\Omega)}^2 \le C\left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)}^{2/3}\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^{4/3} \le C\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^{4/3}.$$

Collecting all estimates yields

$$\begin{aligned}&\left|n_{\text{skew}}\left(\boldsymbol{w},\boldsymbol{w},\boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\boldsymbol{w}^h,\boldsymbol{w}^h,\boldsymbol{\phi}^h\right)\right|\\
&\le \frac{1}{2}\|\boldsymbol{\eta}\|_{L^2(\Omega)}^2 + \frac{1}{4}\|\nabla\cdot\boldsymbol{\eta}\|_{L^2(\Omega)}^2 + C\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^{4/3}\|\nabla\boldsymbol{\eta}\|_{L^3(\Omega)}^2\end{aligned}$$

$$+C \left\| \mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^2 \|\boldsymbol{\eta}\|_{L^6(\Omega)}^2 + \frac{\mu}{4}\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2$$
$$+\left(\frac{3}{4}+\frac{1}{2}\|\nabla\boldsymbol{w}\|_{L^\infty(\Omega)}^2+\frac{1}{4}\|\boldsymbol{w}\|_{L^\infty(\Omega)}^2+\|\nabla\boldsymbol{w}\|_{L^\infty(\Omega)}+\frac{1}{4\mu}\|\boldsymbol{w}\|_{L^\infty(\Omega)}^2\right)\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2.$$

This bound is inserted in the right-hand side of (8.130) giving

$$\frac{1}{2}\frac{d}{dt}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2+\frac{\underline{C}C_S\delta^2}{3}\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^3(\Omega)}^3+\frac{\mu}{2}\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2$$
$$+\frac{1}{2}\left(2\nu+\nu_0(\delta)\right)\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)}^2+\sum_{j=1}^J\frac{\beta}{2}\left\|\boldsymbol{\phi}^h\cdot\hat{\boldsymbol{\tau}}_j\right\|_{L^2(\Gamma_j)}^2$$
$$\le\left[\frac{2}{3\left(\underline{C}C_S\right)^{1/2}\delta}\|\partial_t\boldsymbol{\eta}\|_{V'}^{3/2}+\frac{1}{2}\left(2\nu+\nu_0(\delta)\right)\|\mathbb{D}(\boldsymbol{\eta})\|_{L^2(\Omega)}^2\right.$$
$$+\sum_{j=1}^J\frac{\beta}{2}\left\|\boldsymbol{\eta}\cdot\hat{\boldsymbol{\tau}}_j\right\|_{L^2(\Gamma_j)}^2+\frac{2}{3}\underline{C}^{-1/2}C_S\overline{C}^{3/2}C_L^{3/2}\delta^2\|\mathbb{D}(\boldsymbol{\eta})\|_{L^3(\Omega)}^{3/2}$$
$$+\mu^{-1}\left\|r-q^h\right\|_{L^2(\Omega)}^2+\mu\|\nabla\cdot\boldsymbol{\eta}\|_{L^2(\Omega)}^2+\frac{1}{2}\|\boldsymbol{\eta}\|_{L^2(\Omega)}^2+\frac{1}{4}\|\nabla\cdot\boldsymbol{\eta}\|_{L^2(\Omega)}^2$$
$$\left.+C\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^{4/3}\|\nabla\boldsymbol{\eta}\|_{L^3(\Omega)}^2+C\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^2\|\boldsymbol{\eta}\|_{L^6(\Omega)}^2\right]+\frac{\mu}{4}\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2$$
$$+\left(\frac{3}{4}+\|\nabla\boldsymbol{w}\|_{L^\infty(\Omega)}+\left(\frac{1}{4}+\frac{1}{4\mu}\right)\|\boldsymbol{w}\|_{L^\infty(\Omega)}^2+\frac{1}{2}\|\nabla\boldsymbol{w}\|_{L^\infty(\Omega)}^2\right)\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2.$$

For the application of Gronwall's lemma, Lemma A.55, one needs that

$$\frac{3}{4}+\|\nabla\boldsymbol{w}\|_{L^\infty(\Omega)}+\left(\frac{1}{4}+\frac{1}{4\mu}\right)\|\boldsymbol{w}\|_{L^\infty(\Omega)}^2+\frac{1}{2}\|\nabla\boldsymbol{w}\|_{L^\infty(\Omega)}^2\in L^1(0,T),$$

in other words $\boldsymbol{w}\in L^2\left(0,T;W^{1,\infty}(\Omega)\right)$. The term on the right-hand side of this inequality containing $C_L^{3/2}$ is treated as in the proof of Theorem 8.118. To obtain finally a uniform error estimate, one has also to verify that the terms containing $\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}$ are bounded uniformly in ν. To this end, one gets with Hölder's inequality

$$\int_0^T\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^{4/3}\|\mathbb{D}(\boldsymbol{\eta})\|_{L^3(\Omega)}^2\,dt$$
$$\le\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^{4q/3}\left(0,T;L^3(\Omega)\right)}^{4/3}\|\mathbb{D}(\boldsymbol{\eta})\|_{L^{2q'}\left(0,T;L^3(\Omega)\right)}^2,$$

where $q^{-1} + q'^{-1} = 1$. From the stability estimates (8.125) and (8.126), it follows that one has to take q such that $4q/3 \leq 3$. Accordingly, one can choose $q = 9/4, q' = 9/5$. Inserting this choice gives

$$\int_0^T \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^{4/3} \|\mathbb{D}(\boldsymbol{\eta})\|_{L^3(\Omega)}^2 \, dt$$
$$\leq C \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3\left(0,T;L^3(\Omega)\right)}^{4/3} \|\mathbb{D}(\boldsymbol{\eta})\|_{L^{18/5}\left(0,T;L^3(\Omega)\right)}^2 \leq CC_4(\delta) \|\mathbb{D}(\boldsymbol{\eta})\|_{L^{18/5}\left(0,T;L^3(\Omega)\right)}^2 .$$

Similarly, for the conjugate exponents $q = 3/2,\ q' = 3$, one obtains

$$\begin{aligned}\int_0^T \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3(\Omega)}^2 \|\boldsymbol{\eta}\|_{L^6(\Omega)}^2 \, dt &\leq \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^{2q}\left(0,T;L^3(\Omega)\right)}^2 \|\boldsymbol{\eta}\|_{L^{2q'}\left(0,T;L^6(\Omega)\right)}^2 \\ &\leq \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^3\left(0,T;L^3(\Omega)\right)}^2 \|\boldsymbol{\eta}\|_{L^6\left(0,T;L^6(\Omega)\right)}^2 \\ &\leq C_4(\delta) \|\boldsymbol{\eta}\|_{L^6\left(0,T;L^6(\Omega)\right)}^2 .\end{aligned}$$

The stated error estimate follows now from Gronwall's inequality (A.40) and the triangle inequality as in the proof of Theorem 8.118. ■

Remark 8.121 (Interpretation of the Error Estimates and Numerical Studies) The estimates given in Theorems 8.118 and 8.120 show that the error between the solution of the continuous and the discrete Smagorinsky model in different norms is bounded independently of ν for fixed filter width δ. In this case, the order of convergence is related to the best approximation errors of the finite element spaces in several norms. The best approximation error of V_{div}^h can be estimated by the best approximation error of V^h, see Lemma 3.60 or Remark 3.62 for the $L^2(\Omega)$ norm of the gradient. For instance, considering an error which is squared on the left-hand side of estimate (8.143), then the order of convergence is bounded, e.g., by the best approximation error $\left\|\mathbb{D}\left(\boldsymbol{w} - \tilde{\boldsymbol{w}}^h\right)\right\|_{L^3\left(0,T;L^3(\Omega)\right)}^{3/4}$ in expression (8.144). Numerical studies in John and Layton (2002) and John (2004, Sect. 8.1.8) show that the errors are independent of ν and an even higher order of convergence than predicted by the error bounds. □

Remark 8.122 (Finite Element Analysis of the Fully Discrete Smagorinsky Model) A finite element analysis of the fully discrete Smagorinsky model, with the IMEX Euler scheme as temporal discretization, is presented in Chacón Rebollo and Lewandowski (2014, Chap. 10). Stability estimates are given, the well-posedness of the problem is proved, error estimates are derived, and the asymptotic energy balance is studied. □

Remark 8.123 (Analysis of Finite Element Discretizations for the Stationary Smagorinsky Model) Finite element methods for the stationary Smagorinsky model were studied in Du and Gunzburger (1990). It was proved that the discrete solution converges to the solution of the continuous problem under minimal regularity assumptions on this solution. In addition, an optimal order finite element error estimate for $\left\|\boldsymbol{w}-\boldsymbol{w}^h\right\|_{H^1(\Omega)}$ is given. A comprehensive presentation of a finite element analysis for this model and further results can be found in Chacón Rebollo and Lewandowski (2014, Chap. 9). □

8.3.4 Variants for Reducing Some Drawbacks of the Smagorinsky Model

Remark 8.124 (Drawbacks of the Smagorinsky Model in Numerical Simulations) The advantages of the Smagorinsky model were already mentioned in Remark 8.63: easiness of implementation, robustness, and low costs. However, this model possesses in its application for flow simulations also a number of drawbacks, e.g., see Zang et al. (1993) or Sagaut (2006, pp. 113, 123)

The easiest way, which is in fact quite popular, consists in choosing the Smagorinsky coefficient C_S in (8.66) a priori as a constant. However, it is well known that it is generally not possible to represent the large scales of turbulent flows correctly with a single constant. In addition, a reasonable good choice of C_S depends on many aspects, e.g., the flow problem or the discretization. For a concrete simulation, it is usually not clear what are good values for C_S. Numerical simulations with the Smagorinsky model that can be found in the literature use typically a Smagorinsky constant of size $C_S \in [0.01, 0.1]$, e.g., see Sagaut (2006, p. 124) or Piomelli (1999). For the case of homogeneous isotropic turbulence, in Lilly (1967) the constant $C_S^* \approx 0.17$ in (8.68) ($C_S \approx 0.08$ in (8.67)) was derived, compare also Pope (2000, Sect. 13.4.2). Inappropriate choices of C_S might give very bad computational results, e.g., see Example 8.128.

The Smagorinsky model introduces generally too much viscosity into the flow simulations. This behavior can be observed in particular near solid walls with no-slip boundary conditions.

More drawbacks of the Smagorinsky model arise from the fact that $C_S\delta^2 \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{\mathrm{F}} \geq 0$. Thus, backscatter of energy is prevented. Usually it is $C_S\delta^2 \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{\mathrm{F}} > 0$, even for laminar flows or in subregions where the studied flow field is laminar. Hence, the Smagorinsky model generally does not vanish for laminar flows and thus it introduces also in simulations of such flows unnecessary viscosity. □

Remark 8.125 (Contents of this Section) This section describes two approaches that are used in practice for reducing the drawbacks of the Smagorinsky model. In the dynamic Smagorinsky model, Remark 8.126, the Smagorinsky parameter C_S is computed a posteriori as a function in space and time. To reduce the introduction of model viscosity near solid walls, there is a proposal to decrease C_S, see Remark 8.127, the so-called van Driest damping. □

Remark 8.126 (The Dynamic Smagorinsky Model) The Smagorinsky model (8.66) contains the parameter C_S. As already noted, a good choice of C_S depends on the concrete flow problem and it is in general a priori hardly to achieve. It is even desirable to choose C_S in a different way in different flow regions. In particular, the impact of the Smagorinsky model should be small in subregions with laminar flows, where small values of C_S are preferable. A different approach for restriction the impact of the Smagorinsky model is discussed Remark 8.216.

An approach which determines values for C_S as a function of space and time was proposed in Germano et al. (1991). This proposal was modified in Lilly (1992) to the form presented here. It is called dynamic Smagorinsky model or dynamic subgrid scale model.

The dynamic Smagorinsky model starts with introducing a second filter, a so-called test filter denoted by a hat, with $\hat{\delta} > \delta$. Then, the space-averaged Navier–Stokes equations (8.28) and (8.29)

$$\partial_t \overline{\boldsymbol{u}} - 2\nu\nabla\cdot\mathbb{D}(\overline{\boldsymbol{u}}) + \nabla\cdot\left(\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T\right) + \nabla\cdot\mathbb{T} + \nabla\overline{p} = \overline{\boldsymbol{f}},$$

are filtered once more with the test filter. Assuming that differentiation and filtering commute yields

$$\begin{aligned}\partial_t \widehat{\overline{\boldsymbol{u}}} - 2\nu\nabla\cdot\mathbb{D}\left(\widehat{\overline{\boldsymbol{u}}}\right) + \nabla\cdot\left(\widehat{\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T}\right) + \nabla\cdot\widehat{\mathbb{T}} + \nabla\widehat{\overline{p}} &= \widehat{\overline{\boldsymbol{f}}} \quad \text{in } (0,T]\times\Omega,\\ \nabla\cdot\widehat{\overline{\boldsymbol{u}}} &= 0 \quad \text{in } (0,T]\times\Omega.\end{aligned}$$

A direct calculation gives for

$$\mathbb{K} = \widehat{\overline{\boldsymbol{u}\boldsymbol{u}^T}} - \widehat{\overline{\boldsymbol{u}}}\;\widehat{\overline{\boldsymbol{u}}}^{\,T}$$

that

$$\mathbb{K} - \widehat{\mathbb{T}} = \widehat{\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T} - \widehat{\overline{\boldsymbol{u}}}\;\widehat{\overline{\boldsymbol{u}}}^{\,T}. \tag{8.146}$$

Now, one takes the ansatz of the Smagorinsky model (8.62), (8.66) for both tensors with the same parameter $C_S(t,\boldsymbol{x})$

$$\begin{aligned}\mathbb{T}(t,\boldsymbol{x}) - \frac{\operatorname{trace}(\mathbb{T})}{3}\mathbb{I} &= -C_S(t,\boldsymbol{x})\,\delta^2\,\|\mathbb{D}(\overline{\boldsymbol{u}})\|_{\mathrm{F}}\,\mathbb{D}(\overline{\boldsymbol{u}}),\\ \mathbb{K}(t,\boldsymbol{x}) - \frac{\operatorname{trace}(\mathbb{K})}{3}\mathbb{I} &= -C_S(t,\boldsymbol{x})\,\hat{\delta}^{\,2}\left\|\mathbb{D}\left(\widehat{\overline{\boldsymbol{u}}}\right)\right\|_{\mathrm{F}}\mathbb{D}\left(\widehat{\overline{\boldsymbol{u}}}\right)\end{aligned}$$

Inserting this ansatz in (8.146) yields

$$\begin{aligned}\boldsymbol{0} &= -\widehat{\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T} + \mathbb{K}(t,\boldsymbol{x}) - \mathbb{T}(t,\boldsymbol{x})\\ &= \widehat{\overline{\boldsymbol{u}}}\;\widehat{\overline{\boldsymbol{u}}}^{\,T} + -\widehat{\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T} + \widehat{\overline{\boldsymbol{u}}}\;\widehat{\overline{\boldsymbol{u}}}^{\,T} + \frac{1}{3}\left(\operatorname{trace}(\mathbb{K}) - \widehat{\operatorname{trace}(\mathbb{T})}\right)\mathbb{I} \qquad (8.147)\\ &\quad + \widehat{\left(C_S(t,\boldsymbol{x})\,\delta^2\,\|\mathbb{D}(\overline{\boldsymbol{u}})\|_{\mathrm{F}}\,\mathbb{D}(\overline{\boldsymbol{u}})\right)} - C_S(t,\boldsymbol{x})\,\hat{\delta}^{\,2}\left\|\mathbb{D}\left(\widehat{\overline{\boldsymbol{u}}}\right)\right\|_{\mathrm{F}}\mathbb{D}\left(\widehat{\overline{\boldsymbol{u}}}\right).\end{aligned}$$

From the linearity of the filter (8.25), the linearity of the trace operator, and (8.146), it follows that

$$\begin{aligned}\operatorname{trace}(\mathbb{K}) - \widehat{\operatorname{trace}(\mathbb{T})} &= \operatorname{trace}(\mathbb{K}) - \operatorname{trace}\left(\widehat{\mathbb{T}}\right) = \operatorname{trace}\left(\mathbb{K} - \widehat{\mathbb{T}}\right) \\ &= \operatorname{trace}\left(\widehat{\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T} - \widehat{\overline{\boldsymbol{u}}}\ \widehat{\overline{\boldsymbol{u}}}^{\,T}\right). \end{aligned} \tag{8.148}$$

In order to obtain an equation for $C_S(t,\boldsymbol{x})$, one approximates

$$\widehat{\left(C_S(t,\boldsymbol{x})\,\delta^2\,\|\mathbb{D}(\overline{\boldsymbol{u}})\|_{\mathrm{F}}\,\mathbb{D}(\overline{\boldsymbol{u}})\right)} \approx C_S(t,\boldsymbol{x})\,\delta^2\,\widehat{\left(\|\mathbb{D}(\overline{\boldsymbol{u}})\|_{\mathrm{F}}\,\mathbb{D}(\overline{\boldsymbol{u}})\right)}. \tag{8.149}$$

If C_S depends only on t but not on $\boldsymbol{x}$, one has an equality instead of an approximation. Inserting (8.149) and (8.148) in (8.147) gives

$$\begin{aligned}\boldsymbol{0} \approx & -\widehat{\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T} + \widehat{\overline{\boldsymbol{u}}}\ \widehat{\overline{\boldsymbol{u}}}^{\,T} + \frac{1}{3}\operatorname{trace}\left(\widehat{\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T} - \widehat{\overline{\boldsymbol{u}}}\ \widehat{\overline{\boldsymbol{u}}}^{\,T}\right)\mathbb{I} \\ & + C_S(t,\boldsymbol{x})\left(\delta^2\,\widehat{\left(\|\mathbb{D}(\overline{\boldsymbol{u}})\|_{\mathrm{F}}\,\mathbb{D}(\overline{\boldsymbol{u}})\right)} - \hat{\delta}^{\,2}\left\|\mathbb{D}\left(\widehat{\overline{\boldsymbol{u}}}\right)\right\|_{\mathrm{F}}\mathbb{D}\left(\widehat{\overline{\boldsymbol{u}}}\right)\right) \\ =: & \ \mathbb{L} + C_S\mathbb{M}. \end{aligned} \tag{8.150}$$

Equations for $C_S(t,\boldsymbol{x})$ are obtained by replacing the approximation sign in (8.150) with the equal sign. Then, there are $d(d+1)/2$ equations to determine a scaler value for given t and $\boldsymbol{x}$. Because of the divergence constraint, the traces of the deformation tensors vanish, such that only $d(d+1)/2-1$ equations are linearly independent. In Lilly (1992), it was proposed to determine the parameter $C_S(t,\boldsymbol{x})$ by the least squares method, i.e., to find $C_S(t,\boldsymbol{x})$ such that $\|\mathbb{L} + C_S(t,\boldsymbol{x})\,\mathbb{M}\|_{\mathrm{F}}^2$ is minimized. Evaluating the necessary condition for a minimum gives

$$\begin{aligned} 0 &= \frac{d}{dC_S}\|\mathbb{L} + C_S\mathbb{M}\|_{\mathrm{F}}^2 = \frac{d}{dC_S}\sum_{i,j=0}^{d}\left(\mathbb{L}_{ij} + C_S\mathbb{M}_{ij}\right)^2 \\ &= 2\sum_{i,j=0}^{d}\left(\mathbb{L}_{ij} + C_S\mathbb{M}_{ij}\right)\mathbb{M}_{ij} = 2\sum_{i,j=0}^{d}\mathbb{L}_{ij}\mathbb{M}_{ij} + 2C_S\sum_{i,j=0}^{d}\mathbb{M}_{ij}\mathbb{M}_{ij} \\ &= 2\mathbb{L} : \mathbb{M} + 2C_S\mathbb{M} : \mathbb{M}. \end{aligned}$$

It follows that

$$C_S(t,\boldsymbol{x}) = -\frac{\mathbb{L}:\mathbb{M}}{\mathbb{M}:\mathbb{M}}(t,\boldsymbol{x}). \tag{8.151}$$

In practical computations, the test filter can be applied by solving the space-averaged Navier–Stokes equations on a coarse grid. If the next coarser grid of the current grid is used, then $\hat{\delta} = 2\delta$.

The dynamic subgrid scale model can predict negative values for $C_S(t,\boldsymbol{x})$. In this way, backscatter of energy is possible, in contrast to the Smagorinsky model, see Remark 8.7. However, experience shows that $C_S(t,\boldsymbol{x})$ can vary strongly in space and time and it might contain negative values with a very large amplitude. If the total viscosity $2\nu + \nu_{\mathrm{T}}$ becomes locally non-positive, then one has the discretization of an operator which is not elliptic. The mathematical properties of such an operator are not clear. These properties may strongly destabilize numerical simulations. In practice, the nominator and denominator of (8.151) are averaged, often in time, to compute a smoother function $C_S(t,\boldsymbol{x})$, e.g., see Lesieur (1997, p. 405), Breuer (1998), or Sagaut (2006, p. 139). □

Remark 8.127 (Van Driest Damping) As already mentioned in Remark 8.124, the Smagorinsky model introduces too much viscosity in particular near solid walls. The application of a van Driest damping, see van Driest (1956), is a proposal to reduce this viscosity. The van Driest damping changes the eddy viscosity of the Smagorinsky model (8.66) in the viscous sublayer, see Remark 8.13, to Pope (2000, p. 599)

$$\nu_{\mathrm{T}} = C_S\delta^2\left(1-\exp\left(-\frac{y^+}{A^+}\right)\right)^2 \|\mathbb{D}(\overline{\boldsymbol{u}})\|_{\mathrm{F}}, \quad \text{if } y^+ < 5, \tag{8.152}$$

with $A^+ = 26$. □

Example 8.128 (Turbulent Channel Flow at $Re_\tau = 180$) This example presents some numerical results obtained with the Smagorinsky model for the turbulent channel flow at $\mathrm{Re}_\tau = 180$, see Example D.12.

The domain was triangulated with a grid consisting of $8 \times 16 \times 8$ hexahedra with right angles. In the directions with periodic boundary conditions, the x- and the z-direction, a uniform distance of the grid points was used. The flow exhibits boundary layers at the walls $y = 0$ and $y = 2$, see Fig. D.12. For this reason, it is common to use in the wall-normal y-direction a grid that is refined towards the walls. There are several proposals in the literature for defining the grid points in this direction, e.g., see John and Kindl (2008).

For the results presented in this example, the grid points were set according to

$$y_i = 1 - \cos\left(\frac{i\pi}{N_y}\right), \quad i = 0, \ldots, N_y,$$

where N_y is the number of mesh cells in y-direction.

The simulations were performed with the Q_2/P_1^{disc} pair of finite element spaces. Using this pair results in 25,344 degrees of freedom for the velocity (including Dirichlet nodes) and 4096 pressure degrees of freedom. Compared with most results from the literature, the number of degrees of freedom is quite small. The deformation tensor form of the viscous term, the convective form of the convection term, and the form (8.67) of the Smagorinsky model were used. Using the Smagorinsky model requires the choice of the constant C_S and of the filter width δ. The filter width δ was set to be two times a measure for the local mesh width. However, there is the difficulty of defining the measure for the local mesh width since in the vicinity of the walls there are anisotropic mesh cells that look like thin plates. Possible choices are the diameter h_K, the cubic root of the volume of the mesh cell $h_{K,\mathrm{vol}} = |K|^{1/3}$, and the shortest edge $h_{K,\mathrm{short}}$ of the mesh cell. It holds

$$h_{K,\mathrm{short}} \leq h_{K,\mathrm{vol}} \leq h_K.$$

The larger this measure, the more viscosity is introduced. An appropriate range of values for choosing C_S is known from the literature, e.g., see John and Kindl (2008, 2010b).

The simulations were performed in the time interval $[0, 40]$. Statistics were computed in $[10, 40]$. As temporal discretization, the Crank–Nicolson scheme, see Example 7.49, with the equidistant time step $\Delta t = 0.004$ was used. This length of the time step satisfies condition (D.38). The stopping criterion of the Picard iteration for solving the nonlinear problem in each discrete time was the Euclidean norm of the residual vector to be smaller than 10^{-6}.

A few selected numerical results are presented in Fig. 8.4. From the mean velocity field, it can be seen that $\delta_K = 2h_{K,\mathrm{short}}$ and $C_S = 0.005$ is the only combination that provides a qualitatively correct approximation. Comparing the standard Smagorinsky model and the Smagorinsky model with van Driest damping (8.152) for these values, one can see only little differences. The results obtained with van Driest damping are a little bit more accurate. However, it can be observed that the second order statistics $\mathbb{T}^h_{12,\mathrm{mean}}$ and $u^h_{1,\mathrm{rms}}$ are considerably overpredicted (for $\mathbb{T}^h_{12,\mathrm{mean}}$ with respect to the absolute value).

□

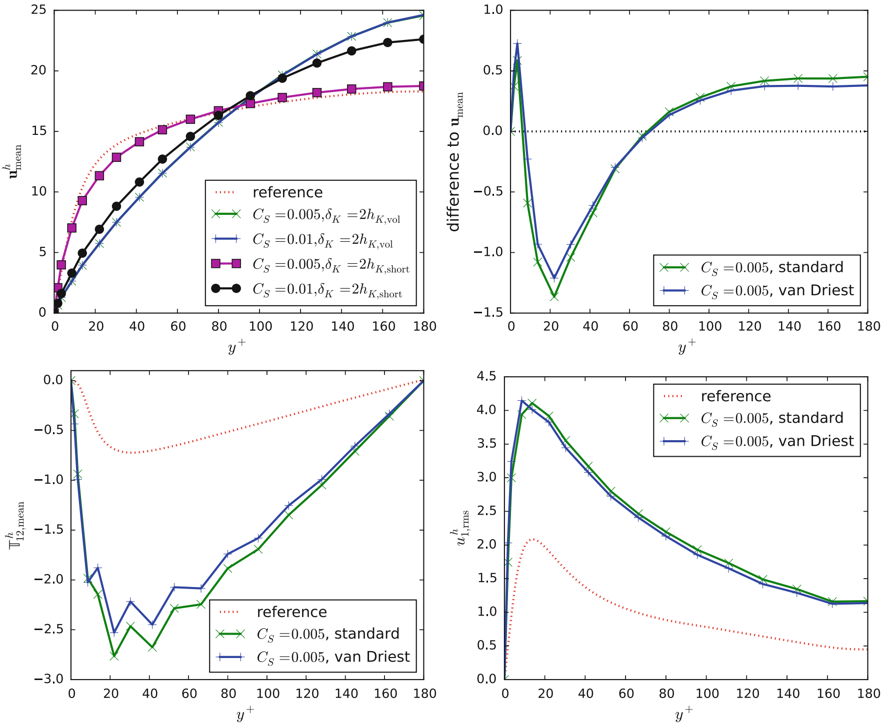

Fig. 8.4 Example 8.128. Mean velocity field for the standard Smagorinsky model with different parameters (*top left, blue line on top of the green line*). The results with $\delta_K = 2h_K$ were even worse than with $\delta_K = 2h_{K,\mathrm{vol}}$. Statistics of interest for the Smagorinsky model with $\delta_K = 2h_{K,\mathrm{short}}$ and $C_S = 0.005$ (*top right, bottom*)

8.4 Large Eddy Simulation: Models Based on Approximations in Wave Number Space

Remark 8.129 (The Basic Approach) Inserting the decomposition (8.11) and applying the linearity (8.25) of the filter yields for filtered nonlinear convective term, which is unknown,

$$\overline{\boldsymbol{u}\boldsymbol{u}^T} = \overline{\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T} + \overline{\overline{\boldsymbol{u}}\,\boldsymbol{u}'^T} + \overline{\boldsymbol{u}'\,\overline{\boldsymbol{u}}^T} + \overline{\boldsymbol{u}'\boldsymbol{u}'^T}. \tag{8.153}$$

The first term in (8.153) is called large scale advective term. It describes the convection of the large eddies driven by themselves. The second and third term are the so-called cross terms describing the interaction of the large scale and subgrid scale components. The last tensor is the subgrid scale (sgs) term that describes how the small eddies extract energy from the flow.

Models that are based on approximations in wave number space consider the terms on the right-hand side of (8.153) separately. Each term is transformed to the Fourier or wave number space and then an approximation is applied, see Sect. 8.4.1. It turns out that with this approach the sgs term is modeled with the zero tensor. Numerical simulations show that this model is not sufficient, e.g., see John (2004, Sect. 10.3.3). Section 8.4.2 presents some proposals for modeling the sgs term. The final models will be presented and discussed in Sect. 8.4.3. □

8.4.1 Modeling of the Large Scale and Cross Terms

Remark 8.130 (Assumptions) Let $\Omega = \mathbb{R}^d$ and let the averaging be performed with the Gaussian filter

$$\overline{\boldsymbol{u}}\ (t,\boldsymbol{x}) = g_{\text{Gauss}} * \boldsymbol{u}\,(t,\boldsymbol{x})\,,$$

where g_{Gauss} is defined in (8.18) and δ is a constant. □

Remark 8.131 (Principal Approach) The model of the large scale and cross terms is obtained in five steps:

1. compute the Fourier transform,
2. replace $\mathcal{F}\,(\boldsymbol{u}')$ by a function of $\mathcal{F}\,(\,\overline{\boldsymbol{u}}\,)$ if necessary,
3. approximate the Fourier transform of the Gaussian filter with a simpler function,
4. neglect all terms that are in a certain sense of higher order in δ,
5. compute the inverse Fourier transform.

There are two approaches in the literature which differ in the third point. The first approach, see Leonard (1975) or Clark et al. (1979), approximates $\mathcal{F}\,(g_{\text{Gauss}})$ by a Taylor polynomial, Remark 8.133, whereas the second approach from Galdi and Layton (2000) uses a rational approximation, see Remark 8.136. The first approach gives the Taylor LES model and the second one the rational LES model. □

Remark 8.132 (Step 1 and 2) Using (A.25), the Fourier transform of the large scale term is

$$\mathcal{F}\left(\overline{\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^{\,T}}\right) = \mathcal{F}\,(g_{\text{Gauss}})\,\mathcal{F}\left(\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^{\,T}\right), \tag{8.154}$$

and the Fourier transforms of the cross terms are

$$\begin{aligned}\mathcal{F}\left(\overline{\overline{\boldsymbol{u}}\,\boldsymbol{u}'^T}\right) &= \mathcal{F}\,(g_{\text{Gauss}})\,\mathcal{F}\left(\overline{\boldsymbol{u}}\,\boldsymbol{u}'^T\right)\\ &= \mathcal{F}\,(g_{\text{Gauss}})\left(\mathcal{F}\,(\,\overline{\boldsymbol{u}}\,) * \mathcal{F}\left(\boldsymbol{u}'\right)^T\right), \qquad (8.155)\\ \mathcal{F}\left(\overline{\boldsymbol{u}'\,\overline{\boldsymbol{u}}^{\,T}}\right) &= \mathcal{F}\,(g_{\text{Gauss}})\left(\mathcal{F}\left(\boldsymbol{u}'\right) * \mathcal{F}\,(\,\overline{\boldsymbol{u}}\,)^T\right).\end{aligned}$$

Since $\mathcal{F}\left(g_{\text{Gauss}}\right) \neq 0$, one obtains

$$\mathcal{F}(\boldsymbol{u}) = \frac{\mathcal{F}\left(g_{\text{Gauss}}\right)\mathcal{F}(\boldsymbol{u})}{\mathcal{F}\left(g_{\text{Gauss}}\right)} = \frac{\mathcal{F}(\overline{\boldsymbol{u}})}{\mathcal{F}\left(g_{\text{Gauss}}\right)}. \tag{8.156}$$

Inserting the decomposition $\boldsymbol{u} = \overline{\boldsymbol{u}} + \boldsymbol{u}'$ in the left-hand side and rearranging terms yields

$$\mathcal{F}\left(\boldsymbol{u}'\right) = \left(\frac{1}{\mathcal{F}\left(g_{\text{Gauss}}\right)} - 1\right)\mathcal{F}(\overline{\boldsymbol{u}}). \tag{8.157}$$

Thus, the Fourier transform of $\boldsymbol{u}'$ can be represented with the Fourier transforms of $\boldsymbol{u}$ and of the Gaussian filter. For this representation, it is important that $\mathcal{F}\left(g_{\text{Gauss}}\right) \neq 0$. Inserting (8.157) in (8.155) gives

$$\begin{aligned}
\mathcal{F}\left(\overline{\overline{\boldsymbol{u}}\,\boldsymbol{u}'^T}\right) &= \mathcal{F}\left(g_{\text{Gauss}}\right)\left(\mathcal{F}(\overline{\boldsymbol{u}}) * \left(\frac{1}{\mathcal{F}\left(g_{\text{Gauss}}\right)} - 1\right)\mathcal{F}(\overline{\boldsymbol{u}})^T\right), \\
\mathcal{F}\left(\overline{\boldsymbol{u}'\,\overline{\boldsymbol{u}}^T}\right) &= \mathcal{F}\left(g_{\text{Gauss}}\right)\left(\left(\frac{1}{\mathcal{F}\left(g_{\text{Gauss}}\right)} - 1\right)\mathcal{F}(\overline{\boldsymbol{u}}) * \mathcal{F}(\overline{\boldsymbol{u}})^T\right).
\end{aligned} \tag{8.158}$$

Although on the right-hand sides $\boldsymbol{u}'$ does not appear, there is no reduction of the complexity of the terms so far since still equality holds. □

Remark 8.133 (Step 3 with Taylor Polynomial Approximation) The Fourier transform of the Gaussian filter is an exponential, see (8.19), and the Taylor series expansion for the exponential has the form

$$e^{ax} = 1 + ax + \mathcal{O}\left(x^2\right).$$

Applying this expansion to $\mathcal{F}\left(g_{\text{Gauss}}\right)$ with respect to δ and for fixed $\boldsymbol{y}$ gives

$$\mathcal{F}\left(g_{\text{Gauss}}\right)(\delta, \boldsymbol{y}) = 1 - \frac{\|\boldsymbol{y}\|_2^2}{24}\delta^2 + \mathcal{O}\left(\delta^4\right). \tag{8.159}$$

Since of $1/\mathcal{F}\left(g_{\text{Gauss}}\right)$ is an exponential as well, one obtains the expansion

$$\frac{1}{\mathcal{F}\left(g_{\text{Gauss}}\right)}(\delta, \boldsymbol{y}) = 1 + \frac{\|\boldsymbol{y}\|_2^2}{24}\delta^2 + \mathcal{O}\left(\delta^4\right). \tag{8.160}$$

Now, $\mathcal{F}\left(g_{\text{Gauss}}\right)$ and $1/\mathcal{F}\left(g_{\text{Gauss}}\right)$ are approximated in (8.154) and (8.158) by quadratic polynomials that are obtained by neglecting the terms of $\mathcal{O}\left(\delta^4\right)$ in (8.159) and (8.160), see Fig. 8.5 for the one-dimensional situation. It can be seen that the polynomial approximation of $\mathcal{F}\left(g_{\text{Gauss}}\right)$ is a good approximation only for small wave numbers and it is completely wrong for high wave numbers. Consequently, the

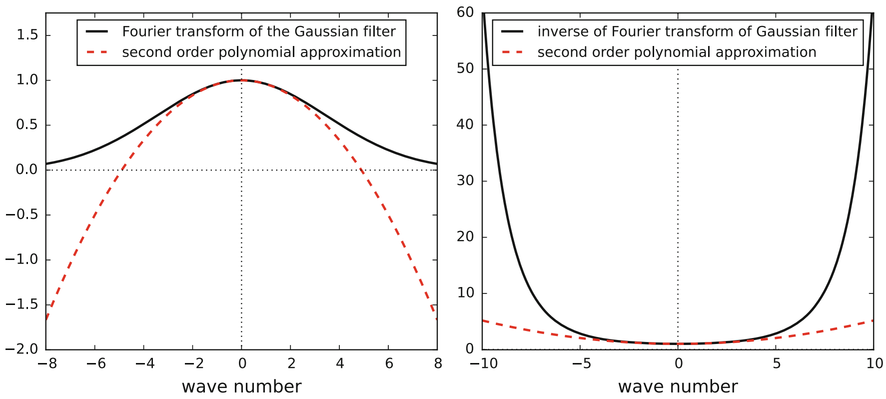

Fig. 8.5 $\mathcal{F}\left(g_{\text{Gauss}}\right)$ (*left*) and $1/\mathcal{F}\left(g_{\text{Gauss}}\right)$ (*right*) with their polynomial approximations, $\delta = 1$

most important property of a filter function, the damping of the high wave number components of $(\boldsymbol{u}, p)$, is not preserved by its Taylor polynomial approximation.

Inserting (8.159) and (8.160) in (8.154) and (8.158) gives

$$\mathcal{F}\left(\overline{\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T}\right) = \left(1 - \frac{\|\boldsymbol{y}\|_2^2}{24}\delta^2 + \mathcal{O}\left(\delta^4\right)\right)\mathcal{F}\left(\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T\right), \tag{8.161}$$

$$\begin{aligned}\mathcal{F}\left(\overline{\overline{\boldsymbol{u}}\,\boldsymbol{u}'^T}\right) &= \left(1 - \frac{\|\boldsymbol{y}\|_2^2}{24}\delta^2 + \mathcal{O}\left(\delta^4\right)\right) \\ &\quad \times \left[\mathcal{F}\left(\overline{\boldsymbol{u}}\right) * \left(\frac{\|\boldsymbol{y}\|_2^2}{24}\delta^2 + \mathcal{O}\left(\delta^4\right)\right)\mathcal{F}\left(\overline{\boldsymbol{u}}\right)^T\right], \\ \mathcal{F}\left(\overline{\boldsymbol{u}'\,\overline{\boldsymbol{u}}^T}\right) &= \left(1 - \frac{\|\boldsymbol{y}\|_2^2}{24}\delta^2 + \mathcal{O}\left(\delta^4\right)\right) \\ &\quad \times \left[\left(\frac{\|\boldsymbol{y}\|_2^2}{24}\delta^2 + \mathcal{O}\left(\delta^4\right)\right)\mathcal{F}\left(\overline{\boldsymbol{u}}\right) * \mathcal{F}\left(\overline{\boldsymbol{u}}\right)^T\right].\end{aligned} \tag{8.162}$$

□

Remark 8.134 (Step 4 with Taylor Polynomial Approximation) Now, the expressions obtained in Step 3 are simplified using properties of the Fourier transform. All terms that have as factor the fourth or a higher power of δ are neglected. However, the other factors of these terms depend on $\overline{\boldsymbol{u}}$ which in turn depends on δ in some unknown way. That means, the neglected terms are only formally of fourth order in

δ. One gets with (A.26) for the large scale advective term from (8.161)

$$\mathcal{F}\left(\overline{\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T}\right) = \mathcal{F}\left(\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T\right) - \frac{\|\boldsymbol{y}\|_2^2}{24}\delta^2\mathcal{F}\left(\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T\right) + \mathcal{O}^{\text{formal}}\left(\delta^4\right)$$
$$= \mathcal{F}\left(\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T\right) + \frac{\delta^2}{24}\mathcal{F}\left(\Delta\left(\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T\right)\right) + \mathcal{O}^{\text{formal}}\left(\delta^4\right).$$

For the first cross term (8.162), one obtains with (A.26) and (A.25)

$$\mathcal{F}\left(\overline{\overline{\boldsymbol{u}}\,\boldsymbol{u}'^T}\right) = \left(1 - \frac{\|\boldsymbol{y}\|_2^2}{24}\delta^2\right)\left(\mathcal{F}\left(\overline{\boldsymbol{u}}\right) * \frac{\|\boldsymbol{y}\|_2^2}{24}\delta^2\mathcal{F}\left(\overline{\boldsymbol{u}}\right)^T\right) + \mathcal{O}^{\text{formal}}\left(\delta^4\right)$$
$$= \left(\mathcal{F}\left(\overline{\boldsymbol{u}}\right) * \frac{\|\boldsymbol{y}\|_2^2}{24}\delta^2\mathcal{F}\left(\overline{\boldsymbol{u}}\right)^T\right) + \mathcal{O}^{\text{formal}}\left(\delta^4\right)$$
$$= -\frac{\delta^2}{24}\left(\mathcal{F}\left(\overline{\boldsymbol{u}}\right) * \mathcal{F}\left(\Delta\overline{\boldsymbol{u}}\right)^T\right) + \mathcal{O}^{\text{formal}}\left(\delta^4\right)$$
$$= -\frac{\delta^2}{24}\mathcal{F}\left(\overline{\boldsymbol{u}}\,\Delta\left(\overline{\boldsymbol{u}}\right)^T\right) + \mathcal{O}^{\text{formal}}\left(\delta^4\right).$$

In the same way, one gets for the other cross term

$$\mathcal{F}\left(\overline{\boldsymbol{u}'\overline{\boldsymbol{u}}^T}\right) = -\frac{\delta^2}{24}\mathcal{F}\left(\Delta\left(\overline{\boldsymbol{u}}\right)\overline{\boldsymbol{u}}^T\right) + \mathcal{O}^{\text{formal}}\left(\delta^4\right).$$

□

Remark 8.135 (Step 5 with Taylor Polynomial Approximation—the Taylor LES Model) The final approximation of the individual terms is computed by applying the inverse Fourier transform

$$\overline{\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T} = \overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T + \frac{\delta^2}{24}\Delta\left(\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T\right) + \mathcal{O}^{\text{formal}}\left(\delta^4\right),$$
$$\overline{\overline{\boldsymbol{u}}\,\boldsymbol{u}'^T} = -\frac{\delta^2}{24}\overline{\boldsymbol{u}}\,\Delta\left(\overline{\boldsymbol{u}}\right)^T + \mathcal{O}^{\text{formal}}\left(\delta^4\right),$$
$$\overline{\boldsymbol{u}'\overline{\boldsymbol{u}}^T} = -\frac{\delta^2}{24}\Delta\left(\overline{\boldsymbol{u}}\right)\overline{\boldsymbol{u}}^T + \mathcal{O}^{\text{formal}}\left(\delta^4\right).$$

In this way, the approximation of the large scale and cross terms of the so-called Taylor LES model reads as follows

$$\overline{\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T} + \overline{\overline{\boldsymbol{u}}\,\boldsymbol{u}'^T} + \overline{\boldsymbol{u}'\overline{\boldsymbol{u}}^T}$$
$$\approx \overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T + \frac{\delta^2}{24}\left(\Delta\left(\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T\right) - \overline{\boldsymbol{u}}\,\Delta\left(\overline{\boldsymbol{u}}\right)^T - \Delta\left(\overline{\boldsymbol{u}}\right)\overline{\boldsymbol{u}}^T\right)$$
$$= \overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T + \frac{\delta^2}{12}\nabla\overline{\boldsymbol{u}}\,\nabla\overline{\boldsymbol{u}}^T, \qquad (8.163)$$

where

$$\Delta\left(\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T\right) - \overline{\boldsymbol{u}}\,\Delta\left(\overline{\boldsymbol{u}}\right)^T - \Delta\left(\overline{\boldsymbol{u}}\right)\overline{\boldsymbol{u}}^T = 2\nabla\overline{\boldsymbol{u}}\,\nabla\overline{\boldsymbol{u}}^T \tag{8.164}$$

was used. Equality (8.164) can be checked by a direct calculation, e.g., considering just an entry of the tensor and a second order derivative gives with the product rule

$$\begin{aligned}\partial_{xx}\left(\overline{u}_i\overline{u}_j\right) &= \partial_x\left(\partial_x\overline{u}_i\overline{u}_j + \overline{u}_i\partial_x\overline{u}_j\right)\\ &= \partial_{xx}\overline{u}_i\overline{u}_j + 2\partial_x\overline{u}_i\partial_x\overline{u}_j + \overline{u}_i\partial_{xx}\overline{u}_j.\end{aligned}$$

□

Remark 8.136 (Step 3 with Rational Approximation) Based on the observation that the Fourier transform of the Gaussian filter is approximated very badly for large wave numbers with the Taylor polynomial approximation, it was proposed in Galdi and Layton (2000) to use a second order rational approximation of the exponential of the form

$$e^{ax} = \frac{1}{1+ax} + \mathcal{O}\left(a^2x^2\right).$$

Applying this subdiagonal Padé approximation to $\mathcal{F}(g_{\text{Gauss}})$ gives

$$\mathcal{F}(g_{\text{Gauss}})(\delta, \boldsymbol{y}) = \frac{1}{1 + \frac{\|\boldsymbol{y}\|_2^2}{24}\delta^2} + \mathcal{O}\left(\delta^4\right) \tag{8.165}$$

and transforming this formula to $1/\mathcal{F}(g_{\text{Gauss}})$ yields

$$\frac{1}{\mathcal{F}(g_{\text{Gauss}})}(\delta, \boldsymbol{y}) = 1 + \frac{\|\boldsymbol{y}\|_2^2}{24}\delta^2 + \mathcal{O}^{\text{formal}}\left(\delta^4\right). \tag{8.166}$$

The last term in (8.166) is actually $\mathcal{O}\left(\delta^4\right)/\mathcal{F}(g_{\text{Gauss}})$ such that it is only formally of fourth order. The rational approximations of $\mathcal{F}(g_{\text{Gauss}})$ and $1/\mathcal{F}(g_{\text{Gauss}})$ are obtained by neglecting all (formal) fourth order terms in (8.165) and (8.166). The behavior of $\mathcal{F}(g_{\text{Gauss}})$ for high wave numbers is much better approximated than with the Taylor polynomial, see Fig. 8.6 for a one-dimensional sketch. The approximation of $1/\mathcal{F}(g_{\text{Gauss}})$ is the same as in the Taylor polynomial case. □

Remark 8.137 (Steps 4 and 5 with Rational Approximation) The derivation of the model continues now in the same way as in the Taylor polynomial case. Inserting (8.165) and (8.166) in (8.154) and (8.158), simplifying the arising terms using properties of the Fourier transform like (A.25), (A.26), (A.28), neglecting all terms that are formally of fourth order with respect to δ and applying the inverse

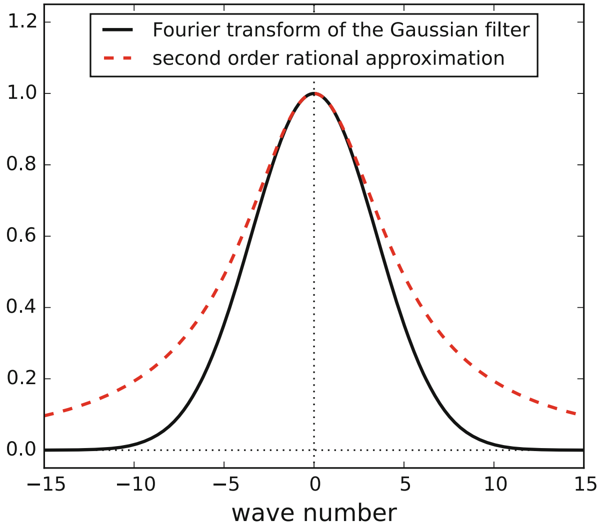

Fig. 8.6 $\mathcal{F}(g_{\text{Gauss}})$ with its second order rational approximations, $\delta = 1$

Fourier transform yields

$$\begin{aligned}
\overline{\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T} &= \left(I - \frac{\delta^2}{24}\Delta\right)^{-1} \left(\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T\right) + \mathcal{O}^{\text{formal}}\left(\delta^4\right), \qquad (8.167)\\
\overline{\overline{\boldsymbol{u}}\,\boldsymbol{u}'^T} &= -\frac{\delta^2}{24}\left(I - \frac{\delta^2}{24}\Delta\right)^{-1} \left(\overline{\boldsymbol{u}}\,\Delta\left(\overline{\boldsymbol{u}}\right)^T\right) + \mathcal{O}^{\text{formal}}\left(\delta^4\right),\\
\overline{\boldsymbol{u}'\,\overline{\boldsymbol{u}}^T} &= -\frac{\delta^2}{24}\left(I - \frac{\delta^2}{24}\Delta\right)^{-1} \left(\Delta\left(\overline{\boldsymbol{u}}\right)\overline{\boldsymbol{u}}^T\right) + \mathcal{O}^{\text{formal}}\left(\delta^4\right).
\end{aligned}$$

□

Remark 8.138 (The (Second Order) Rational LES Model) The approximation of the large scale and the cross terms for the so-called (second order) rational LES model has the form, using (8.164) in the derivation,

$$\begin{aligned}
&\overline{\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T} + \overline{\overline{\boldsymbol{u}}\,\boldsymbol{u}'^T} + \overline{\boldsymbol{u}'\,\overline{\boldsymbol{u}}^T}\\
&\approx \left(I - \frac{\delta^2}{24}\Delta\right)^{-1}\left[\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T - \frac{\delta^2}{24}\left(\overline{\boldsymbol{u}}\,\Delta\left(\overline{\boldsymbol{u}}\right)^T + \Delta\left(\overline{\boldsymbol{u}}\right)\overline{\boldsymbol{u}}^T\right)\right]\\
&= \left(I - \frac{\delta^2}{24}\Delta\right)^{-1}\left[\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T - \frac{\delta^2}{24}\Delta\left(\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T\right) + \frac{\delta^2}{12}\nabla\overline{\boldsymbol{u}}\,\nabla\overline{\boldsymbol{u}}^T\right]\\
&= \overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T + \frac{\delta^2}{12}\left(I - \frac{\delta^2}{24}\Delta\right)^{-1}\nabla\overline{\boldsymbol{u}}\,\nabla\overline{\boldsymbol{u}}^T. \qquad (8.168)
\end{aligned}$$

In Galdi and Layton (2000, Formula (2.10)), there is a misprint (minus sign instead of plus sign). □

Remark 8.139 (On the Operator in the Rational LES Model) The operator $\left(I - \delta^2/24\Delta\right)^{-1}$ describes an elliptic, second order problem which has to be solved

$$-\frac{\delta^2}{24}\Delta\overline{\mathbb{U}} + \overline{\mathbb{U}} = \nabla\overline{\boldsymbol{u}}\nabla\overline{\boldsymbol{u}}^T. \tag{8.169}$$

This problem is a Helmholtz equation. It is called differential filter in turbulence modeling. In connection with the rational LES model, it is usually called auxiliary problem. Differential filters are used also in the definition of approximate deconvolution models (ADM), see (8.178), the Leray-α model turbulence model (8.191), and the Navier–Stokes-α model (8.215).

In the analysis of the incompressible Navier–Stokes equations, operators of form $\left(I - \delta^2/24\Delta\right)^{-1}$ are called Yosida approximation of the identity, e.g., see Sohr (2001, Sect. II.3.4). Such operators approximate functions from $L^2(\Omega)$ by more regular functions.

Some properties of the differential filter will be discussed in Sect. 8.5, starting with Remark 8.153, and of the Galerkin finite element discretization of the differential filter in Sect. 8.6, starting from Remark 8.176. From (8.167), it can be already observed that the operator $\left(I - \delta^2/24\Delta\right)^{-1}$ is an approximation of the convolution with the Gaussian filter, compare also Remark 8.141. □

Remark 8.140 (The Differential Filter in a Bounded Domain) If Ω is a bounded domain, which is usually the case in computations, the differential filter has to be equipped with boundary conditions on $\partial\Omega$. In Galdi and Layton (2000), it is proposed to use homogeneous Neumann boundary conditions. The only exception is the case that periodic boundary conditions are prescribed for the flow problem at some part of the boundary. Then, the differential filter is equipped also with periodic boundary conditions at those parts of the boundary.

This kind of boundary conditions were applied in simulations, e.g., in Iliescu et al. (2003) and John et al. (2010). In these simulations, the use of homogeneous Neumann boundary conditions gave generally reasonable results. □

Remark 8.141 (The Differential Filter is an Approximation of the Convolution) A direct calculation, using the rational approximation (8.165) of $\mathcal{F}(g_{\text{Gauss}})$ and the property (A.28) of the Fourier transform, gives

$$\begin{aligned}\mathcal{F}(g_{\text{Gauss}} * \boldsymbol{u}) = \mathcal{F}(g_{\text{Gauss}})\,\mathcal{F}(\boldsymbol{u}) &\approx \frac{1}{1 + \frac{\|\boldsymbol{y}\|_2^2}{24}\delta^2}\mathcal{F}(\boldsymbol{u}) \\ &= \mathcal{F}\left(\left(I - \frac{\delta^2}{24}\Delta\right)^{-1}\boldsymbol{u}\right),\end{aligned}$$

such that the application of the inverse Fourier transform yields

$$g_{\text{Gauss}} * \boldsymbol{u} \approx \left(I - \frac{\delta^2}{24}\Delta\right)^{-1} \boldsymbol{u}. \tag{8.170}$$

Thus, the differential filter is an approximation of the convolution operator with the Gaussian filter.

Relation (8.170) suggests that the rational LES model can be defined with a convolution instead of the auxiliary problem

$$\overline{\overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T} + \overline{\overline{\boldsymbol{u}}\,\boldsymbol{u}'^T} + \overline{\boldsymbol{u}'\,\overline{\boldsymbol{u}}^T} \approx \overline{\boldsymbol{u}}\,\overline{\boldsymbol{u}}^T + \frac{\delta^2}{12} g_{\text{Gauss}} * \left(\nabla \overline{\boldsymbol{u}}\, \nabla \overline{\boldsymbol{u}}^T\right).$$

This model is called rational LES model with convolution. However, it is practically not used, partly because of the overhead for approximating the convolution operator, see John (2004, Sect. 7.8). □

Remark 8.142 (A Fourth Order Rational Approximation) The second order polynomial and rational approximations model the sgs term $\overline{\boldsymbol{u}'\boldsymbol{u}'^T}$ by $\mathbf{0}$, see Remark 8.144. Using a fourth order rational approximation, one gets a non-trivial model for the sgs term, see John (2004, Sect. 4.2.3). However, the arising approximation of the sgs tensor involves the solution of a fourth order partial differential equation and it involves some higher order terms that are hard to approximate. The only attempt to explore this model, however only from the point of view of analysis, can be found in Berselli and Iliescu (2003). □

8.4.2 Models for the Subgrid Scale Term

Remark 8.143 (On the SGS Term) The sgs term $\overline{\boldsymbol{u}'\boldsymbol{u}'^T}$ is considered to possess a great influence on the formation of turbulence. Thus, its modeling is of great importance. □

Remark 8.144 (The Second Order Fourier Transform Approach) If the sgs term is modeled with the second order approaches that were used for the large scale term and the cross terms, Sect. 8.4.1, one obtains with the second order polynomial approximation of the Fourier transform of the Gaussian filter

$$\overline{\boldsymbol{u}'\boldsymbol{u}'^T} = \frac{\delta^4}{576}\left(\Delta \overline{\boldsymbol{u}}\, \Delta \overline{\boldsymbol{u}}^T\right) + \mathcal{O}^{\text{formal}}\left(\delta^6\right)$$

and with the second order rational approximation

$$\overline{\boldsymbol{u}'\boldsymbol{u}'^T} = \frac{\delta^4}{576}\left(I - \frac{\delta^2}{24}\Delta\right)^{-1} \left(\Delta \overline{\boldsymbol{u}}\, \Delta \overline{\boldsymbol{u}}^T\right) + \mathcal{O}^{\text{formal}}\left(\delta^6\right).$$

Both approximations are formally of fourth order in δ and therefore they will be neglected in these approaches. That means, one obtains the approximation

$$\overline{\boldsymbol{u}'\boldsymbol{u}'^T} \approx \mathbf{0},$$

which proves to be not robust for long time simulations, see John (2004, Sect. 10.3.3) for a numerical example. □

Remark 8.145 (The Smagorinsky Model) A popular approach consists in using the Smagorinsky model (8.67) for approximating the sgs term. However, with the Fourier space approach, Remark 8.144, one obtains that the sgs term is formally of fourth order in δ whereas the Smagorinsky model is formally of second order in δ. There is a contradiction, at least formally. Recall that in the formal higher order terms there is some dependency on δ which is not known.

In the rational LES models, the Smagorinsky model is used only as a model of the sgs term. Thus, its influence should be kept smaller than in computations with the pure Smagorinsky model, which can be achieved by choosing a smaller constant C_S for the rational LES model than for the pure Smagorinsky model. □

Remark 8.146 (Models Based on Physical Arguments) In Iliescu and Layton (1998), several eddy viscosity models for modeling the sgs term based on physical arguments were derived. One of these models has the form

$$\nu_{\mathrm{T}} = C_S\delta \left\| \overline{\boldsymbol{u}} - g_{\mathrm{Gauss}} * \overline{\boldsymbol{u}} \right\|_2 . \tag{8.171}$$

Using (A.25) gives

$$\begin{aligned} \overline{\boldsymbol{u}} - g_{\mathrm{Gauss}} * \overline{\boldsymbol{u}} &= \mathcal{F}^{-1}\left(\mathcal{F}\left(\overline{\boldsymbol{u}} - g_{\mathrm{Gauss}} * \overline{\boldsymbol{u}}\right)\right) \\ &= \mathcal{F}^{-1}\left(\mathcal{F}\left(\overline{\boldsymbol{u}}\right) - \mathcal{F}\left(g_{\mathrm{Gauss}}\right)\mathcal{F}\left(\overline{\boldsymbol{u}}\right)\right) \\ &= \mathcal{F}^{-1}\left(\left(1 - \mathcal{F}\left(g_{\mathrm{Gauss}}\right)\right)\mathcal{F}\left(\overline{\boldsymbol{u}}\right)\right). \end{aligned}$$

For the polynomial approximation (8.159), it is obvious that the term in parentheses is $\mathcal{O}\left(\delta^2\right)$. Using the rational approximation (8.165) yields

$$1 - \mathcal{F}\left(g_{\mathrm{Gauss}}\right) = \frac{1 + \frac{\|\boldsymbol{y}\|_2^2}{24}\delta^2 - 1}{1 + \frac{\|\boldsymbol{y}\|_2^2}{24}\delta^2} + \mathcal{O}\left(\delta^4\right) = \mathcal{O}\left(\delta^2\right).$$

Altogether, model (8.171) is of third order in δ.

In numerical simulation, model (8.171) is usually applied in such a way that the convolution operator is approximated by a second order partial differential operator, see Remark 8.141, i.e.,

$$\nu_{\mathrm{T}} = C_S\delta \left\| \overline{\boldsymbol{u}} - \left(I - \frac{\delta^2}{24}\Delta\right)^{-1} \overline{\boldsymbol{u}} \right\|_2 . \tag{8.172}$$

□

8.4.3 The Resulting Models

Remark 8.147 (Models Based on Approximations in Wave Number Space) The models derived in Sects. 8.4.1 and 8.4.2 can be written concisely in the following strong form

$$
\begin{aligned}
\partial_t \boldsymbol{w} - \nabla \cdot ((2\nu + \nu_{\mathrm{T}}) \mathbb{D}(\boldsymbol{w})) + (\boldsymbol{w} \cdot \nabla) \boldsymbol{w} & \\
+ \nabla r + \nabla \cdot \frac{\delta^2}{12} \left(A_{\text{wave}} \left(\nabla \boldsymbol{w} \nabla \boldsymbol{w}^T \right) \right) &= \overline{\boldsymbol{f}} \quad \text{in } (0,T] \times \Omega, \\
\nabla \cdot \boldsymbol{w} &= 0 \quad \text{in } (0,T] \times \Omega, \\
\boldsymbol{w}(0,\cdot) &= \boldsymbol{w}_0 \text{ in } \Omega,
\end{aligned}
\tag{8.173}
$$

where $(\boldsymbol{w}, r)$ should be an approximation to $(\overline{\boldsymbol{u}}, \overline{p})$. The operator A_{wave} depends on the approximation of the Fourier transform of the Gaussian filter:

- $A_{\text{wave}} = I$ for the Taylor LES model (8.163),
- $A_{\text{wave}} = \left(I - \delta^2/(24)\,\Delta\right)^{-1}$ for the second order rational LES model with auxiliary problem (8.168),
- $A_{\text{wave}} = g_{\text{Gauss}} *$ for the second order rational LES model with convolution (8.171).

Possible choices for the model of the sgs term ν_{T} are the Smagorinsky model (8.66) and the Iliescu–Layton model (8.172).

In a bounded domain, usually the same boundary conditions are applied for $\boldsymbol{w}$ as they are prescribed for $\boldsymbol{u}$. The differential filter is equipped with homogeneous Neumann boundary conditions, see Remark 8.140. □

Remark 8.148 (Analytical Results)

- For the Taylor LES model with Smagorinsky sgs term, the existence, uniqueness, and stability of a weak solution for all times was proved in Coletti (1997) under the assumption that $\nu_{\mathrm{T}} \geq \delta^2/6$. This condition means that the Smagorinsky term dominates the Taylor LES model. This relation is not correct since the Taylor LES model is a model for the large scale and cross term, which are $\mathcal{O}^{\text{formal}}\left(\delta^2\right)$, whereas the sgs term is $\mathcal{O}^{\text{formal}}\left(\delta^4\right)$. Under a similar condition, a finite element error analysis was performed in Iliescu et al. (2002).
- The existence and uniqueness of a solution of the rational LES model was studied in Berselli et al. (2002). This model was considered with auxiliary problem but without subgrid scale term, $\nu_{\mathrm{T}} = 0$ in (8.173). In addition, the case of a space-periodic setting was investigated, i.e., $\Omega = (0, L)^3$ and periodic boundary conditions on $\partial\Omega$. Periodic boundary conditions are applied in the auxiliary problem as well. The existence and uniqueness of a solution in an appropriate function space for small time intervals $T = \mathcal{O}\left(\delta^4\right)$ could be proved. It was already mentioned in Remark 8.144 that the rational LES model without model for the sgs term is not robust in long time simulations. The proof uses the

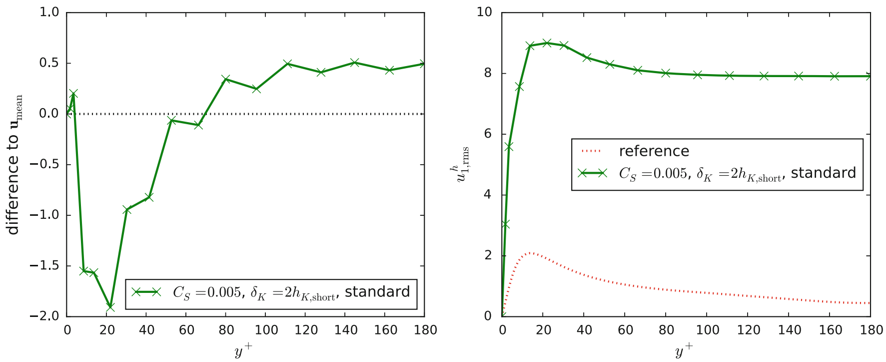

Fig. 8.7 Example 8.149. Rational LES model (8.173) with $A_{\text{wave}} = \left(I - \delta^2/\left(24\right)\Delta\right)^{-1}$, $\delta_K = 2h_{K,\text{short}}$, and with the Smagorinsky model as model of the sgs term

Galerkin method in a similar way as described in Sect. 7.1. Further analytical results can be found in Barbato et al. (2007).

A finite element error analysis of the rational LES model is not available.

□

Example 8.149 (Turbulent Channel Flow at $Re_\tau = 180$) This benchmark problem is described in Example D.12. The setup of the simulations was exactly as for the Smagorinsky LES model, compare Example 8.128. The auxiliary problem was equipped with homogeneous Neumann boundary conditions.

Results for the rational LES model (8.173) with auxiliary problem, i.e., with $A_{\text{wave}} = \left(I - \delta^2/\left(24\right)\Delta\right)^{-1}$, and with the Smagorinsky model as model for the sgs term are presented in Fig. 8.7. Actually, one would expect that with the rational LES model, the constant in the Smagorinsky model should be smaller than for the Smagorinsky LES model, since the Smagorinsky model is used to model a term that is $\mathcal{O}^{\text{formal}}\left(\delta^4\right)$. The most reasonable result was obtained in fact for $C_S = 0.005$. Simulations with smaller values of C_S and with the van Driest damping (8.152) and $C_S = 0.005$ blew up. With $C_S = 0.005$, the mean velocity field is approximated reasonably well. However, only a very bad prediction of the second statistic $u^h_{1,\text{rms}}$ is obtained.

Similar conclusions with respect to the first and second order statistics can be drawn from the results for the rational LES model with the Iliescu–Layton model for ν_{T}, compare Fig. 8.8.

As mentioned in Example D.12, it is advised to incorporate contributions of the actually used turbulence models in the computation of the second order statistics. In the performed simulations, such contributions were not included. Perhaps, their inclusion would improve the results somewhat. □

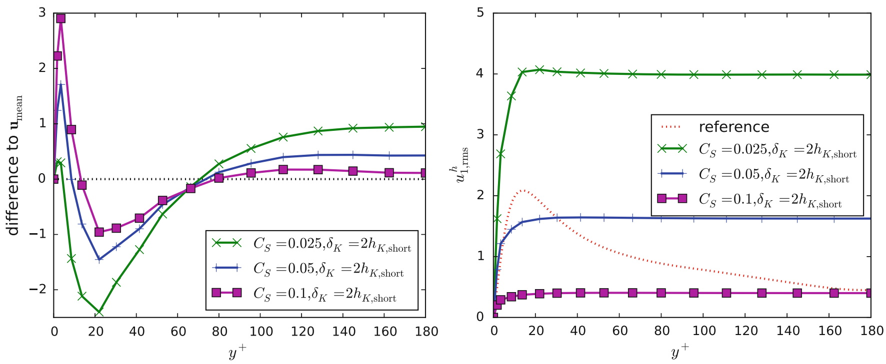

Fig. 8.8 Example 8.149. Rational LES model (8.173) with $A_{\text{wave}} = \left(I - \delta^2/(24)\,\Delta\right)^{-1}$, $\delta_K = 2h_{K,\text{short}}$, and with the Iliescu–Layton model as model for the sgs term

8.5 Large Eddy Simulation: Approximate Deconvolution Models (ADMs)

Remark 8.150 (Basic Idea) Approximate deconvolution models (ADM) were introduced in Stolz and Adams (1999) and Stolz et al. (2001). Consider the space-averaged Navier–Stokes equations

$$\begin{aligned} \partial_t \overline{\boldsymbol{u}} - 2\nu\nabla\cdot\mathbb{D}(\overline{\boldsymbol{u}}) + \nabla\cdot\left(\overline{\boldsymbol{u}\boldsymbol{u}^T}\right) + \nabla\overline{p} = \overline{\boldsymbol{f}} \text{ in } (0,T]\times\mathbb{R}^d, \\ \nabla\cdot\overline{\boldsymbol{u}} = 0 \text{ in } (0,T]\times\mathbb{R}^d, \end{aligned} \tag{8.174}$$

see (8.28). As already discussed in Remark 8.32, the momentum equation in (8.174) is not yet an equation for $(\overline{\boldsymbol{u}}, \overline{p})$ since the nonlinear convective term still depends on $\boldsymbol{u}$.

Often, the filter is defined by a convolution with an appropriate filter function. For this reason, the filter operator is denoted by G_{conv}, i.e., $\overline{\boldsymbol{u}} = G_{\text{conv}}(\boldsymbol{u})$. If the filter operator is invertible, then $\boldsymbol{u} = G_{\text{deconv}}(\overline{\boldsymbol{u}})$, where $G_{\text{deconv}} = G_{\text{conv}}^{-1}$ is the inverse filter operator or the deconvolution operator. Then, the momentum equation of (8.174) becomes

$$\partial_t \overline{\boldsymbol{u}} - 2\nu\nabla\cdot\mathbb{D}(\overline{\boldsymbol{u}}) + \nabla\cdot\left(\overline{G_{\text{deconv}}(\overline{\boldsymbol{u}})G_{\text{deconv}}(\overline{\boldsymbol{u}})^T}\right) + \nabla\overline{p} = \overline{\boldsymbol{f}}, \tag{8.175}$$

which is an equation for $(\overline{\boldsymbol{u}}, \overline{p})$, involving the filter operator and its inverse.

Even if G_{deconv} exists and would be efficiently computable, (8.175) does not define a turbulence model since the nonlinear term contains still all scales of the flow and it has the same complexity as the nonlinear term of the Navier–Stokes equations. The proposal of Stolz and Adams (1999) and Stolz et al. (2001) consists in replacing in (8.175) the deconvolution operator by an approximation of the deconvolution (or of the inverse filter). Denoting the approximate deconvolution of order N by

$G_{\text{deconv},N}$ gives the momentum equation

$$\partial_t \overline{\boldsymbol{u}} - 2\nu\nabla\cdot\mathbb{D}(\overline{\boldsymbol{u}}) + \nabla\cdot\left(\overline{G_{\text{deconv},N}(\overline{\boldsymbol{u}})G_{\text{deconv},N}(\overline{\boldsymbol{u}})^T}\right) + \nabla\overline{p} = \overline{\boldsymbol{f}}, \qquad (8.176)$$

where the order will be specified more precisely in Remark 8.156. Equation (8.176) is an equation for $(\overline{\boldsymbol{u}}, \overline{p})$ with the known and computable filter operator and approximate deconvolution operator.

Since ADMs compute an approximation of $(\overline{\boldsymbol{u}}, \overline{p})$, the solution will be denoted also in this case by $(\boldsymbol{w}, r)$. Hence, an ADM has the form

$$\begin{aligned} \partial_t \boldsymbol{w} - 2\nu\nabla\cdot\mathbb{D}(\boldsymbol{w}) & \\ +\nabla\cdot\left(\overline{G_{\text{deconv},N}(\boldsymbol{w})G_{\text{deconv},N}(\boldsymbol{w})^T}\right) + \nabla r = \overline{\boldsymbol{f}} & \text{ in } (0,T]\times\mathbb{R}^d, \\ \nabla\cdot\boldsymbol{w} = 0 & \text{ in } (0,T]\times\mathbb{R}^d. \end{aligned} \qquad (8.177)$$

□

Remark 8.151 (Mathematical Literature) There are a number of papers on the analysis and numerical analysis of ADMs. An overview of the available results can be found in the monograph Layton and Rebholz (2012).

Note that the space-average Navier–Stokes equations (8.174), which are the basis of ADMs, were derived with the assumption of the commutation of the filter operator and the differential operators, see Remark 8.28. As it was discussed in Sect. 8.2, this assumption is usually not satisfied for filters that are defined by convolution. An exception is the case of space-periodic boundary conditions. In ADMs, usually the differential filter, see (8.169) or (8.178) below, is used. The commutation error of the differential filter will be discussed in Remark 8.154 and Example 8.155. Surveying the mathematical literature of ADMs, one finds that exclusively space-periodic boundary conditions are considered. As already mentioned in Remark 8.2, this kind of boundary conditions is not in the focus here. Since a survey of mathematical results is available, Layton and Rebholz (2012), here only some selected topics of ADMs will be presented. □

Example 8.152 (Inverse Deconvolution Operators) In general, the inverse operator of the filter operator does not need to exist. However, for some examples, one can give the explicit inverse operator.

- For the Gaussian filter, a representation of the deconvolution operator is obtained by applying the inverse Fourier transform to (8.156)

$$\boldsymbol{u} = G_{\text{deconv}}(\overline{\boldsymbol{u}}) = \mathcal{F}^{-1}(\mathcal{F}(\boldsymbol{u})) = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(\overline{\boldsymbol{u}})}{\mathcal{F}(g_{\text{Gauss}})}\right).$$

- Let the filter be defined by the differential filter

$$-\delta^2\Delta\overline{\boldsymbol{u}} + \overline{\boldsymbol{u}} = \boldsymbol{u} \quad \text{in } \Omega, \qquad (8.178)$$

which was already introduced within the framework of the rational LES model in (8.169). This equation has to be equipped with boundary conditions. A possible choice is $\overline{\boldsymbol{u}} = \boldsymbol{u}$ on $\partial\Omega$. Then, it holds

$$\overline{\boldsymbol{u}} = G_{\text{conv}}(\boldsymbol{u}) = \left(-\delta^2\Delta + I\right)^{-1}\boldsymbol{u},$$
$$\boldsymbol{u} = G_{\text{deconv}}\left(\overline{\boldsymbol{u}}\right) = \left(-\delta^2\Delta + I\right)\left(\overline{\boldsymbol{u}}\right).$$

□

Remark 8.153 (On the Boundary Conditions for the Differential Filter) Assuming that $\boldsymbol{u}$ possesses Dirichlet boundary conditions, then a straightforward choice consists in using the same conditions for $\overline{\boldsymbol{u}}$. If these conditions are no-slip conditions, then one finds from the definition of the differential filter that not only $\overline{\boldsymbol{u}} = \boldsymbol{0}$ at the boundary but also $\Delta\overline{\boldsymbol{u}} = \boldsymbol{0}$ at the boundary. Thus, $\overline{\boldsymbol{u}}$ is near the boundary (in the layer) close to a harmonic function, e.g., a linear function, which probably does not correctly reflect the physical behavior of the large scales. □

Remark 8.154 (The Commutation of Filtering and Spatial Derivatives for the Differential Filter) Let $\boldsymbol{u}$ and $\overline{\boldsymbol{u}}$ be sufficiently smooth, then it follows by differentiating the strong form of the filter equation and using the Theorem of Schwarz that

$$\nabla\boldsymbol{u} = \nabla\left(\overline{\boldsymbol{u}} - \delta^2\Delta\overline{\boldsymbol{u}}\right) = \nabla\overline{\boldsymbol{u}} - \delta^2\Delta\left(\nabla\overline{\boldsymbol{u}}\right).$$

On the other hand, the differential filter, extended to tensor-valued functions, for $\nabla\boldsymbol{u}$ is by definition

$$\nabla\boldsymbol{u} = \overline{\nabla\boldsymbol{u}} - \delta^2\Delta\left(\overline{\nabla\boldsymbol{u}}\right).$$

Thus, $\nabla\overline{\boldsymbol{u}}$ and $\overline{\nabla\boldsymbol{u}}$ satisfy the same elliptic partial differential equation with the same right-hand side. If they would satisfy the same boundary condition, then they are identical and filtering and commutation commute.

In the situation of space-periodic boundary conditions for $\boldsymbol{u}$, also space-periodic boundary conditions for $\overline{\boldsymbol{u}}$ will be prescribed. Then, also $\nabla\overline{\boldsymbol{u}}$ and $\overline{\nabla\boldsymbol{u}}$ are space-periodic, i.e., they satisfy the same boundary condition, and thus differentiation and filtering commute.

In the case that Ω is a bounded domain, Dirichlet boundary conditions are prescribed for $\boldsymbol{u}$, and the same boundary conditions are prescribed for each filtered function as for the corresponding unfiltered function, then the boundary conditions are generally not identical, see Example 8.155. □

Example 8.155 (Commutation Error for the Differential Filter in a Bounded Domain) Consider $u(x) = \sin(x) + \delta^2\sin(x)$ in $(0, \pi)$. Then it is $u(0) = u(\pi) = 0$. Using the same boundary conditions for the filter, the differential filter (8.178) is equipped with $\overline{u}(0) = \overline{u}(\pi) = 0$, such that $\overline{u}(x) = \sin(x)$.

On the one hand, it is $\overline{u}'(x) = \cos(x)$ and consequently $\overline{u}'(0) = 1$, $\overline{u}'(\pi) = -1$. For defining the filter of $u'(x) = \cos(x) + \delta^2\cos(x)$, one takes the same boundary conditions as for the unfiltered function, i.e., $\overline{u'}(0) = u'(0) = 1 + \delta^2$ and $\overline{u'}(\pi) = u'(\pi) = -1 - \delta^2$. These are not the same boundary vales as for $\overline{u}'(x)$ and consequently $\overline{u'}(x) \neq \overline{u}'(x)$. □

Remark 8.156 (Van Cittert Approximate Deconvolution) The probably most popular approach for defining an approximate deconvolution is the van Cittert approximate deconvolution, van Cittert (1931).

Let $G_{\text{conv}} : L^2(\Omega) \to L^2(\Omega)$ be a convolution operator, i.e., $G_{\text{conv}}(\boldsymbol{u}) = \overline{\boldsymbol{u}}$. Consider the fixed point equation

$$\boldsymbol{u} = \boldsymbol{u} + (\overline{\boldsymbol{u}} - G_{\text{conv}}(\boldsymbol{u})) . \tag{8.179}$$

The N-th order van Cittert deconvolution $G_{\text{deconv},N}(\overline{\boldsymbol{u}})$ is defined by applying a fixed point iteration to (8.179) for approximating $\boldsymbol{u}$, starting with $G_{\text{deconv},0}(\overline{\boldsymbol{u}}) = \overline{\boldsymbol{u}}$ and performing N steps:

$$\begin{aligned} G_{\text{deconv},0}(\overline{\boldsymbol{u}}) &= \overline{\boldsymbol{u}}, \\ G_{\text{deconv},n}(\overline{\boldsymbol{u}}) &= G_{\text{deconv},n-1}(\overline{\boldsymbol{u}}) + (\overline{\boldsymbol{u}} - G_{\text{conv}}(G_{\text{deconv},n-1}(\overline{\boldsymbol{u}}))), \end{aligned} \tag{8.180}$$

$n = 1, \ldots, N$. □

Remark 8.157 (Properties of the Van Cittert Approximate Deconvolution)

- The first members of the family of van Cittert approximate deconvolutions can be derived in a straightforward way from (8.180)

$$\begin{aligned} G_{\text{deconv},0}(\overline{\boldsymbol{u}}) &= \overline{\boldsymbol{u}}, \\ G_{\text{deconv},1}(\overline{\boldsymbol{u}}) &= G_{\text{deconv},0}(\overline{\boldsymbol{u}}) + (\overline{\boldsymbol{u}} - G_{\text{conv}}(G_{\text{deconv},0}(\overline{\boldsymbol{u}}))) \\ &= \overline{\boldsymbol{u}} + (\overline{\boldsymbol{u}} - G_{\text{conv}}(\overline{\boldsymbol{u}})) = 2\overline{\boldsymbol{u}} - \overline{\overline{\boldsymbol{u}}}, \\ G_{\text{deconv},2}(\overline{\boldsymbol{u}}) &= 2\overline{\boldsymbol{u}} - \overline{\overline{\boldsymbol{u}}} + \left(\overline{\boldsymbol{u}} - G_{\text{conv}}\left(2\overline{\boldsymbol{u}} - \overline{\overline{\boldsymbol{u}}}\right)\right) \\ &= 2\overline{\boldsymbol{u}} - \overline{\overline{\boldsymbol{u}}} + \overline{\boldsymbol{u}} - 2\overline{\overline{\boldsymbol{u}}} + \overline{\overline{\overline{\boldsymbol{u}}}} \\ &= 3\overline{\boldsymbol{u}} - 3\overline{\overline{\boldsymbol{u}}} + \overline{\overline{\overline{\boldsymbol{u}}}}. \end{aligned}$$

  It can be seen that the approximate deconvolution of order N is defined by a sum that involves terms with multiple (at most N) applications of the filter to the function $\boldsymbol{u}$.
- Rewriting the fixed point iteration (8.180), one finds the recursion

$$G_{\text{deconv},n}(\overline{\boldsymbol{u}}) = (I - G_{\text{conv}})\, G_{\text{deconv},n-1}(\overline{\boldsymbol{u}}) + \overline{\boldsymbol{u}}, \quad n = 1, \ldots, N.$$

  Straightforward calculations, using $G_{\text{deconv},0}(\overline{\boldsymbol{u}}) = \overline{\boldsymbol{u}}$, give

$$\begin{aligned} G_{\text{deconv},0}(\overline{\boldsymbol{u}}) &= (I - G_{\text{conv}})^0\, G_{\text{deconv},0}(\overline{\boldsymbol{u}}) = \overline{\boldsymbol{u}}, \\ G_{\text{deconv},1}(\overline{\boldsymbol{u}}) &= (I - G_{\text{conv}})\, \overline{\boldsymbol{u}} + (I - G_{\text{conv}})^0\, \overline{\boldsymbol{u}}, \end{aligned}$$

$$
\begin{aligned}
G_{\text{deconv},2}\left(\overline{\boldsymbol{u}}\right) &= \left(I - G_{\text{conv}}\right) G_{\text{deconv},1}\left(\overline{\boldsymbol{u}}\right) + \left(I - G_{\text{conv}}\right)^0 G_{\text{deconv},0}\left(\overline{\boldsymbol{u}}\right) \\
&= \left(I - G_{\text{conv}}\right)^2 \overline{\boldsymbol{u}} + \left(I - G_{\text{conv}}\right) \overline{\boldsymbol{u}} + \left(I - G_{\text{conv}}\right)^0 \overline{\boldsymbol{u}},
\end{aligned}
$$

from what one finds by induction that

$$
G_{\text{deconv},N}\left(\overline{\boldsymbol{u}}\right) = \sum_{n=0}^{N} \left(I - G_{\text{conv}}\right)^n \left(\overline{\boldsymbol{u}}\right), \quad 0 \le N < \infty. \tag{8.181}
$$

□

Lemma 8.158 (Representation of the Error of the Approximate Deconvolution) *Consider the space-periodic case and the differential filter. Then, it is for any* $v \in L^2(\Omega)$

$$
v - G_{\text{deconv},N} G_{\text{conv}} v = (-1)^{N+1} \delta^{2N+2} \left(\Delta G_{\text{conv}}\right)^{N+1} v. \tag{8.182}
$$

Proof Using the binomial theorem, one finds that

$$
G_{\text{conv}} \left(I - G_{\text{conv}}\right)^n = G_{\text{conv}} + n G_{\text{conv}}^2 + \ldots + G_{\text{conv}}^{n+1} = \left(I - G_{\text{conv}}\right)^n G_{\text{conv}}, \quad n \ge 0.
$$

From rewriting (8.181) term by term, it follows that $G_{\text{conv}} G_{\text{deconv},N} = G_{\text{deconv},N} G_{\text{conv}}$. Applying the representation (8.181) of the van Cittert approximate deconvolution, one obtains, using the canceling of terms in telescopic sums,

$$
\begin{aligned}
G_{\text{deconv},N} G_{\text{conv}} = G_{\text{conv}} G_{\text{deconv},N} &= G_{\text{conv}} \sum_{n=0}^{N} \left(I - G_{\text{conv}}\right)^n \\
&= \sum_{n=0}^{N} \left(I - G_{\text{conv}}\right)^n - \left(I - G_{\text{conv}}\right) \sum_{n=0}^{N} \left(I - G_{\text{conv}}\right)^n \\
&= \sum_{n=0}^{N} \left(I - G_{\text{conv}}\right)^n - \sum_{n=1}^{N+1} \left(I - G_{\text{conv}}\right)^n \\
&= I - \left(I - G_{\text{conv}}\right)^{N+1}.
\end{aligned}
$$

It follows that

$$
v - G_{\text{deconv},N} G_{\text{conv}} v = \left(I - G_{\text{conv}}\right)^{N+1} v. \tag{8.183}
$$

Since for the differential filter it is

$$
(-\delta^2 \Delta + I) G_{\text{conv}} = I,
$$

one gets

$$I - G_{\text{conv}} = -\delta^2 \Delta G_{\text{conv}}.$$

Inserting this expression in (8.183) gives (8.182). ■

Lemma 8.159 (Selfadjointness of the Filter Operator) *Let the filter be given by convolution* (8.12) *with a symmetric filter function or by the differential filter* (8.178)*, where either homogeneous Dirichlet boundary conditions are used or the space-periodic case is considered. Then it holds*

$$(\overline{u}, v) = (u, \overline{v}) \quad \forall\, u, v \in L^2(\Omega). \tag{8.184}$$

Proof Filtering with the convolution has to be considered in $\mathbb{R}^d$. Then, one obtains with (8.12), the application of Fubini's theorem, and the symmetry of the filter function

$$\begin{aligned}
(\overline{u}, v) &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} g_{\text{fil}}\,(\boldsymbol{x} - \boldsymbol{z})\, u\,(\boldsymbol{z})\ d\boldsymbol{z} \right) v(\boldsymbol{x})\ d\boldsymbol{x} \\
&= \int_{\mathbb{R}^d} u\,(\boldsymbol{z}) \left(\int_{\mathbb{R}^d} g_{\text{fil}}\,(\boldsymbol{x} - \boldsymbol{z})\, v(\boldsymbol{x})\ d\boldsymbol{x} \right)\ d\boldsymbol{z} \\
&= \int_{\mathbb{R}^d} u\,(\boldsymbol{z}) \left(\int_{\mathbb{R}^d} g_{\text{fil}}\,(\boldsymbol{z} - \boldsymbol{x})\, v(\boldsymbol{x})\ d\boldsymbol{x} \right)\ d\boldsymbol{z} = (u, \overline{v}).
\end{aligned}$$

With the assumed boundary conditions, the variational form of the differential filter of u is given by

$$(u, w) = (\overline{u}, w) + \delta^2 (\nabla \overline{u}, \nabla w),$$

where the test functions are from $H_0^1(\Omega)$ or the corresponding space with space-periodic boundary conditions. For $v \in L^2(\Omega)$, the function $\overline{v}$ is sufficiently smooth such that it can be used as test function, since it also obeys the correct boundary conditions. Choosing $w = \overline{v}$ gives

$$(u, \overline{v}) = (\overline{u}, \overline{v}) + \delta^2 (\nabla \overline{u}, \nabla \overline{v}).$$

With the same arguments, one can use $\overline{u}$ as test function. Using the definition of the differential filter for v with $w = \overline{u}$ yields

$$(\overline{u}, v) = (v, \overline{u}) = (\overline{v}, \overline{u}) + \delta^2 (\nabla \overline{v}, \nabla \overline{u}),$$

which proves the statement for the differential filter. ■

Remark 8.160 (On the Analysis of ADMs) As already mentioned at the beginning of this section, the analysis for ADMs is performed exclusively for the space-periodic case. To give a flavor of applying the properties of the deconvolution operator in this case, the proof of a stability estimate for the lowest order ADM will be presented below. A short survey of available results will be given in Remark 8.163. □

Remark 8.161 (The Lowest Order ADM) The weak formulation of the lowest order ADM in the space-periodic case with the differential filter has the form: Find $(\boldsymbol{w}, r)$ such that

$$\int_0^T \Big[(\partial_t \boldsymbol{w}, \boldsymbol{v}) + (\nu \nabla \boldsymbol{w}, \nabla \boldsymbol{v}) + \left(\nabla \cdot \left(\overline{\boldsymbol{w}\boldsymbol{w}^T}\right), \boldsymbol{v}\right) - (\nabla \cdot \boldsymbol{v}, r) + (\nabla \cdot \boldsymbol{w}, q) \Big]\, dt = \int_0^T \left(\overline{\boldsymbol{f}}, \boldsymbol{v}\right)\, dt \quad \forall\, (\boldsymbol{v}, q), \tag{8.185}$$

where all functions belong to appropriate spaces. □

Lemma 8.162 (Stability Estimate (Energy Equality) for the Zeroth Order ADM) *Let* $\boldsymbol{w}$ *be a sufficiently regular solution of* (8.185)*, then it holds for all* $T > 0$

$$\begin{aligned}
&\frac{1}{2} \|\boldsymbol{w}(T)\|_{L^2(\Omega)}^2 + \frac{\delta^2}{2} \|\nabla \boldsymbol{w}(T)\|_{L^2(\Omega)}^2 + \nu \|\nabla \boldsymbol{w}\|_{L^2(0,T;L^2(\Omega))}^2 \\
&\quad + \delta^2 \nu \|\Delta \boldsymbol{w}\|_{L^2(0,T;L^2(\Omega))}^2 \\
&= \frac{1}{2} \|\boldsymbol{w}(0)\|_{L^2(\Omega)}^2 + \frac{\delta^2}{2} \|\nabla \boldsymbol{w}(0)\|_{L^2(\Omega)}^2 + \int_0^T (\boldsymbol{f}, \boldsymbol{w})\, dt.
\end{aligned} \tag{8.186}$$

Proof As always, one has to choose the test function such that the nonlinear convective term vanishes. Using the commutation of filtering and differentiation in the space-periodic case, (8.184), and (2.29) gives

$$\begin{aligned}
\left(\nabla \cdot \left(\overline{\boldsymbol{w}\boldsymbol{w}^T}\right), \boldsymbol{v}\right) &= \left(\overline{\nabla \cdot (\boldsymbol{w}\boldsymbol{w}^T)}, \boldsymbol{v}\right) = \left(\nabla \cdot \left(\boldsymbol{w}\boldsymbol{w}^T\right), \overline{\boldsymbol{v}}\right) \\
&= \left((\boldsymbol{w} \cdot \nabla)\, \boldsymbol{w}, \overline{\boldsymbol{v}}\right) = n_{\mathrm{conv}}\left(\boldsymbol{w}, \boldsymbol{w}, \overline{\boldsymbol{v}}\right).
\end{aligned}$$

Hence, one has to choose $\overline{\boldsymbol{v}} = \boldsymbol{w}$ which is equivalent to choosing $\boldsymbol{v} = \left(-\delta^2 \Delta + I\right) \boldsymbol{w}$. One has for sufficiently smooth functions

$$\begin{aligned}
\nabla \cdot \Delta \boldsymbol{w} &= \partial_x \left(\partial_{xx} w_1 + \partial_{yy} w_1 + \partial_{zz} w_1\right) \\
&\quad + \partial_y \left(\partial_{xx} w_2 + \partial_{yy} w_2 + \partial_{zz} w_2\right) \\
&\quad + \partial_z \left(\partial_{xx} w_3 + \partial_{yy} w_3 + \partial_{zz} w_3\right) \\
&= \left(\partial_{xx} + \partial_{yy} + \partial_{zz}\right) \left(\partial_x w_1 + \partial_y w_2 + \partial_z w_3\right) = \Delta \left(\nabla \cdot \boldsymbol{w}\right).
\end{aligned} \tag{8.187}$$

Since $\nabla \cdot \boldsymbol{w} = 0$, see (8.177), it follows that $\nabla \cdot \Delta \boldsymbol{w} = 0$. Inserting this test function in (8.185) yields

$$\int_0^T \left[\left(\partial_t \boldsymbol{w}, \left(-\delta^2 \Delta + I\right) \boldsymbol{w}\right) + \left(\nu \nabla \boldsymbol{w}, \nabla \left(\left(-\delta^2 \Delta + I\right) \boldsymbol{w}\right)\right)\right] dt$$
$$= \int_0^T \left(\overline{\boldsymbol{f}}, \left(-\delta^2 \Delta + I\right) \boldsymbol{w}\right) dt.$$

Applying integration by parts, where the boundary integrals cancel due to the space-periodic conditions, and using (8.184) gives

$$\int_0^T \left[(\partial_t \boldsymbol{w}, \boldsymbol{w}) + \delta^2 (\partial_t \nabla \boldsymbol{w}, \nabla \boldsymbol{w}) + (\nu \nabla \boldsymbol{w}, \nabla \boldsymbol{w}) + \delta^2 (\nu \Delta \boldsymbol{w}, \Delta \boldsymbol{w})\right] dt$$
$$= \int_0^T \left(\boldsymbol{f}, \overline{\left(-\delta^2 \Delta + I\right) \boldsymbol{w}}\right) dt = \int_0^T (\boldsymbol{f}, \boldsymbol{w})\ dt. \tag{8.188}$$

Applying (7.13) leads to

$$\int_0^T \left[\frac{d}{dt} \frac{1}{2} \|\boldsymbol{w}\|^2_{L^2(\Omega)} + \frac{\delta^2}{2} \frac{d}{dt} \|\nabla \boldsymbol{w}\|^2_{L^2(\Omega)} + (\nu \nabla \boldsymbol{w}, \nabla \boldsymbol{w}) + \delta^2 (\nu \Delta \boldsymbol{w}, \Delta \boldsymbol{w})\right] dt$$
$$= \int_0^T (\boldsymbol{f}, \boldsymbol{w})\ dt.$$

Integration on $(0, T)$ gives finally (8.186). ■

Remark 8.163 (Short Survey of Analytical Results for ADMs) If not stated otherwise, the results mentioned in this remark are for the space-periodic case, the differential filter, and the van Cittert approximate deconvolution.

- The existence of a weak solution of the zeroth order ADM was proved in Layton and Lewandowski (2003) and its uniqueness was shown in Layton and Lewandowski (2006), see also Layton and Rebholz (2012).
- ADMs of order N were studied in Dunca and Epshteyn (2006). In this paper, the existence and uniqueness of a weak solution was proved. The proof of the existence uses the Galerkin method, which was introduced in Sect. 7.1. The existence and uniqueness results obtained indicate that ADMs are less complex than the three-dimensional Navier–Stokes equations, compare Remark 8.22.
- Another main topic in Dunca and Epshteyn (2006) is an estimate of the modeling error, i.e., the error between $\overline{\boldsymbol{u}}$ and $\boldsymbol{w}$. It was shown that for sufficiently smooth solutions it holds that $\|\overline{\boldsymbol{u}} - \boldsymbol{w}\|_{L^\infty(0,T;L^2(\Omega))}$ and $\|\nabla (\overline{\boldsymbol{u}} - \boldsymbol{w})\|_{L^2(0,T;L^2(\Omega))}$ are $\mathcal{O}\left(\delta^{N+1}\right)$. The proof uses the error representation (8.182).
- In Stanculescu (2008), a general approximate deconvolution operator $G_{\text{deconv},N}$ is considered in combination with the differential filter. Conditions on $G_{\text{deconv},N}$ were derived that ensure the existence and uniqueness of a weak solution, which

guarantee regularity of the weak solution, and which ensure an energy equality. The conditions on $G_{\mathrm{deconv},N}$ are

- $G_{\mathrm{deconv},N}$ is a bounded linear operator on $L^2(\Omega)$,
- $G_{\mathrm{deconv},N}$ is selfadjoint and positive definite,
- $G_{\mathrm{deconv},N}$ commutes with differentiation.

It is pointed out that there are deconvolution operators besides the van Cittert deconvolution that satisfy these conditions and also deconvolution operators which do not.

- The convergence of an ADM for fixed δ for $N \to \infty$ was studied in Berselli and Lewandowski (2012). It was shown that a subsequence of solutions $\{(\boldsymbol{w}_N, r_N)\}_{N=1}^{\infty}$ converges to a solution, in an appropriate sense, of the space-averaged Navier–Stokes equations (8.28).

□

Remark 8.164 (Finite Element Error Analysis of ADMs)

- A numerical scheme for the zeroth order approximate deconvolution model is studied in Manica and Merdan (2007). In this paper, the situation of a bounded domain and no-slip boundary conditions is considered. Hence, differentiation and filtering do not commute and the momentum balance has the form

$$\partial_t \boldsymbol{w} - \nu \Delta \boldsymbol{w} + \overline{\nabla \cdot (\boldsymbol{w}\, \boldsymbol{w}^T)} + \overline{\nabla q} = \overline{\boldsymbol{f}}.$$

To handle the nonlinear convective term, it is proposed to use the same kind of test function as in the proof of Lemma 8.162, namely $\left(-\delta^2 \Delta + I\right) \tilde{\boldsymbol{v}}$. Then, this term can be rewritten in the form $n_{\mathrm{skew}}(\boldsymbol{w}, \boldsymbol{w}, \tilde{\boldsymbol{v}})$ and it vanishes for $\tilde{\boldsymbol{v}} = \boldsymbol{w}$. However, the use of this kind of test functions leads to a fourth order term, compare the term $\delta^2 (\nu \Delta \boldsymbol{w}, \Delta \boldsymbol{w})$ in (8.188). A mixed finite element formulation is considered to handle this situation, which is analyzed in the usual way. A main tool of the analysis is a modified Stokes projection. The optimal choice of δ comes from the properties of the modified Stokes projection and it is $\delta = \mathcal{O}(h)$.
- A so-called reduced ADM was proposed and analyzed in Galvin et al. (2014). The derivation of this model starts with defining the Helmholtz filter with an additional incompressibility constraint. Then, the filtered variable is inserted in the momentum equation of the Navier–Stokes equations and the arising fourth order terms are approximated by second order terms using the approximate deconvolution operator. This approach gives an additional viscous term of the form $-\delta^2 \Delta(\partial_t \boldsymbol{w})$. A similar approach was used for deriving a reduced Navier–Stokes-α model, see Example 8.208. For the reduced ADM, a Crank–Nicolson IMEX scheme was analyzed, unconditional stability of the solution was proved, and error estimates with respect to the velocity solution of the continuous reduced ADM and to the filtered velocity of the Navier–Stokes equations were shown.

□

8.6 The Leray-α Model

Remark 8.165 (Motivation) In Leray (1934b), the existence of a weak solution of the Navier–Stokes equations (8.1) (in this paper called turbulent solution) was proved by considering a sequence of simplified problems, where the simplification consisted in replacing in the nonlinear term of the Navier–Stokes equations the convection field by a smooth or regularized velocity field, see Remark 7.10. The case $\Omega = \mathbb{R}^3$ was considered and the regularization was defined by a convolution with a filter function. Then, the behavior was studied for the filter width tending to zero. Based on this idea from the analysis of the Navier–Stokes equations, a turbulence model can be proposed, the so-called Leray-α model. □

Remark 8.166 (The Regularization Operator) Since the numerical calculation of convolution operators is expensive and in the case of a bounded domain one has to consider a cut-off of the domain of integration, see Remark 8.153, the regularization operator in the Leray-α model is usually the differential filter. It was already noted in Remark 8.141 that the differential filter is an approximation of the convolution with the Gaussian filter function. □

Remark 8.167 (Transforming an Abstract Regularized Equation into an Equation for the Large Scales) Following Geurts and Holm (2003, 2006), a regularization model can be expressed similarly to the basic form of Eq. (8.28) of LES with the sgs tensor given in (8.29). Consider a model of the form

$$\begin{aligned} \partial_t \boldsymbol{w} - \nu \Delta \boldsymbol{w} + (\overline{\boldsymbol{w}} \cdot \nabla) \boldsymbol{w} + \nabla r &= \boldsymbol{f}, \\ \nabla \cdot \boldsymbol{w} &= 0, \\ R \overline{\boldsymbol{w}} &= \boldsymbol{w}, \end{aligned} \tag{8.189}$$

where R is some linear regularization operator that is invertible. It will be assumed that the regularization operator and its inverse commute with derivatives.

Using the commutation property of the inverse operator and its linearity yields

$$\nabla \cdot \overline{\boldsymbol{w}} = \nabla \cdot \left(R^{-1} \boldsymbol{w}\right) = R^{-1} \left(\nabla \cdot \boldsymbol{w}\right) = 0. \tag{8.190}$$

In Geurts and Holm (2003, 2006), the model was equipped from the beginning with the constraint (8.190).

Next, one can write the convective term in the form $\nabla \cdot \left(\boldsymbol{w} \overline{\boldsymbol{w}}^T\right)$, see (2.28). Replacing now in the first equation $\boldsymbol{w}$ with $R\overline{\boldsymbol{w}}$ gives

$$\partial_t \left(R\overline{\boldsymbol{w}}\right) - \nu\Delta \left(R\overline{\boldsymbol{w}}\right) + \nabla \cdot \left(\left(R\overline{\boldsymbol{w}}\right) \overline{\boldsymbol{w}}^T\right) + \nabla r = \boldsymbol{f}.$$

With the assumption of the commutation of the regularization operator with all derivatives, one obtains

$$\begin{aligned} &R\left(\partial_t \overline{\boldsymbol{w}}\right) - R\left(\nu \Delta \overline{\boldsymbol{w}}\right) + R\left(\nabla \cdot \left(\overline{\boldsymbol{w}}\, \overline{\boldsymbol{w}}^T\right)\right) + \nabla r \\ &= \boldsymbol{f} - \left(\nabla \cdot \left(\left(R\overline{\boldsymbol{w}}\right) \overline{\boldsymbol{w}}^T\right) - R\left(\nabla \cdot \left(\overline{\boldsymbol{w}}\, \overline{\boldsymbol{w}}^T\right)\right)\right). \end{aligned}$$

The application of the inverse operator, defining $\overline{r} = R^{-1}r$, $\overline{\boldsymbol{f}} = R^{-1}\boldsymbol{f}$, and using again the commutation property and the linearity of R^{-1} leads to

$$\begin{aligned}
&\partial_t \overline{\boldsymbol{w}} - \nu\Delta\overline{\boldsymbol{w}} + \nabla\cdot\left(\overline{\boldsymbol{w}}\,\overline{\boldsymbol{w}}^T\right) + \nabla\overline{r} \\
&= \overline{\boldsymbol{f}} - \left(\nabla\cdot\left(R^{-1}\left(\left(R\overline{\boldsymbol{w}}\right)\overline{\boldsymbol{w}}^T\right)\right) - \nabla\cdot\left(\overline{\boldsymbol{w}}\,\overline{\boldsymbol{w}}^T\right)\right) \\
&= \overline{\boldsymbol{f}} - \nabla\cdot\left(\overline{\boldsymbol{w}\overline{\boldsymbol{w}}^T} - \overline{\boldsymbol{w}}\,\overline{\boldsymbol{w}}^T\right).
\end{aligned}$$

In this way, the regularized model (8.189) is written in a similar form as the abstract LES model (8.28), where the sgs stress tensor $\mathbb{T}$ from (8.29) is replaced by an asymmetric tensor which contains in its first term already the filtered velocity field. Note that the asymmetry contradicts the first property of models for the sgs stress tensor given in Remark 8.64. □

8.6.1 The Continuous Problem

Remark 8.168 (The Leray-α Model) Let $\Omega = \mathbb{R}^d$, $d \in \{2, 3\}$, be a bounded domain with Lipschitz boundary Γ and let $\alpha > 0$ be a constant. Then, the Leray-α model with homogeneous boundary conditions for the velocity is given by

$$\begin{aligned}
\partial_t \boldsymbol{w} - \nu\Delta\boldsymbol{w} + (\overline{\boldsymbol{w}}\cdot\nabla)\boldsymbol{w} + \nabla r &= \boldsymbol{f} && \text{in } (0,T]\times\Omega, \\
\nabla\cdot\boldsymbol{w} &= 0 && \text{in } (0,T]\times\Omega, \\
\boldsymbol{w} &= \boldsymbol{0} && \text{on } (0,T]\times\Gamma, \\
\boldsymbol{w}(0,\cdot) &= \boldsymbol{u}_0 && \text{in } \Omega, \\
-\alpha^2\Delta\overline{\boldsymbol{w}} + \overline{\boldsymbol{w}} &= \boldsymbol{w} && \text{in } (0,T]\times\Omega, \\
\overline{\boldsymbol{w}} &= \boldsymbol{0} && \text{on } (0,T]\times\Gamma, \\
\int_\Omega r\,d\boldsymbol{x} &= 0 && \text{in } (0,T].
\end{aligned} \tag{8.191}$$

The smoothed or regularized velocity field is obtained by the solution of three scalar Helmholtz equations (differential filter).

In the analysis, the case of periodic boundary conditions, see Remark 2.31, is considered, i.e., Ω is a cube and $\boldsymbol{w}$ and $\overline{\boldsymbol{w}}$ are equipped with space-periodic boundary conditions. □

Remark 8.169 (The Divergence of $\overline{\boldsymbol{w}}$) Assuming sufficient smoothness of $\overline{\boldsymbol{w}}$, taking the divergence of the differential filter, and using (8.187) gives

$$-\alpha^2\Delta\left(\nabla\cdot\overline{\boldsymbol{w}}\right) + \nabla\cdot\overline{\boldsymbol{w}} = \boldsymbol{0} \quad \text{in } (0,T]\times\Omega. \tag{8.192}$$

Equation (8.192) is a Helmholtz equation for $\nabla\cdot\overline{\boldsymbol{w}}$ with homogeneous right-hand side. Thus, its solution depends only on the boundary conditions.

In the case of a bounded domain, boundary values for $\nabla \cdot \overline{\boldsymbol{w}}$ cannot be derived from (8.191). Hence, it is not clear whether or not $\overline{\boldsymbol{w}}$ is divergence-free. From the practical point of view, this issue might not be that important since finite element velocities are usually not (weakly) divergence-free and thus there is no reason why the (discrete) regularization should be divergence-free.

For periodic boundary conditions, it follows from the periodicity of $\overline{\boldsymbol{w}}$ that also $\nabla \cdot \overline{\boldsymbol{w}}$ possesses periodic boundary conditions. Since the Helmholtz problem with periodic boundary conditions has a unique solution, which can be proved, e.g., by an easy extension of Kreiss and Lorenz (2004, Lemma 9.1.2), one gets that $\nabla \cdot \overline{\boldsymbol{w}} = 0$ is this solution. Hence, $\overline{\boldsymbol{w}}$ is divergence-free.

In Geurts and Holm (2003, 2006), where the constraint $\nabla \cdot \overline{\boldsymbol{w}} = 0$ was used, an example was considered that possesses periodic boundary conditions in two directions and a free-slip condition in the third direction. However, the third direction was not of importance for the turbulent character of the considered flow.

□

Theorem 8.170 (Existence and Uniqueness of a Solution in the Space-Periodic Case) *Let $\Omega = (0, 2\pi L)^3$, $L > 0$, and (8.191) be equipped with periodic boundary conditions*

$$
\begin{aligned}
\partial_t \boldsymbol{w} - \nu \Delta \boldsymbol{w} + (\overline{\boldsymbol{w}} \cdot \nabla)\boldsymbol{w} + \nabla r &= \boldsymbol{f} && \text{in } (0,T] \times \Omega, \\
\nabla \cdot \boldsymbol{w} &= 0 && \text{in } (0,T] \times \Omega, \\
(\boldsymbol{w}, r) \text{ is periodic} & && \text{on } (0,T] \times \Gamma, \\
\boldsymbol{w}(0,\cdot) &= \boldsymbol{u}_0 && \text{in } \Omega, \\
-\alpha^2 \Delta \overline{\boldsymbol{w}} + \overline{\boldsymbol{w}} &= \boldsymbol{w} && \text{in } (0,T] \times \Omega, \\
\overline{\boldsymbol{w}} \text{ is periodic} & && \text{on } (0,T] \times \Gamma, \\
\int_\Omega r \, d\boldsymbol{x} &= 0 && \text{in } (0,T].
\end{aligned}
\tag{8.193}
$$

Let

$$
H_{\mathrm{div,per}}(\Omega) = \left\{ \boldsymbol{v} \ : \ \boldsymbol{v} \in L^2(\Omega), \nabla \cdot \boldsymbol{v} = 0, \boldsymbol{v} \text{ is periodic in} \Omega, \int_\Omega \boldsymbol{v} \, d\boldsymbol{x} = 0 \right\},
$$

and $V = H_{\mathrm{div,per}}(\Omega) \cap H^1(\Omega)$.

If $\boldsymbol{f} \in H_{\mathrm{div,per}}(\Omega)$ and $\boldsymbol{u}_0 \in V$, then (8.193) has a unique weak solution which is even a strong solution in $(0,T)$. That means, $\boldsymbol{w}$ satisfies

$$
\frac{d}{dt}(\boldsymbol{w}, \boldsymbol{v}) + \nu (\nabla \boldsymbol{w}, \nabla \boldsymbol{v}) + n(\overline{\boldsymbol{w}}, \boldsymbol{w}, \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v}) \quad \forall \, \boldsymbol{v} \in V,
$$

where $\overline{\boldsymbol{w}} = \left(I - \alpha^2 \Delta\right)^{-1} \boldsymbol{w}$ and

$$
\boldsymbol{w} \in C([0,T]; V) \cap L^2\left((0,T); V \cap H^2(\Omega)\right), \quad \partial_t \boldsymbol{w} = L^2\left((0,T); H_{\mathrm{div,per}}(\Omega)\right).
$$

Proof This theorem is stated in Cheskidov et al. (2005).

The proof for the case $\Omega = \mathbb{R}^3$ follows from the analysis of Leray (1934b). For the proof in the periodic case, it is mentioned in Cheskidov et al. (2005) that similar arguments can be applied as in Foias et al. (2002). ■

Remark 8.171 (On Theorem 8.170) The existence and uniqueness of a weak solution of (8.193) can be proved even with weaker regularity of the data $\boldsymbol{f}$ and $\boldsymbol{u}_0$, see Cheskidov et al. (2005) for the concrete statement. □

Remark 8.172 (Existence of an Attractor) It is shown in Cheskidov et al. (2005) that with the regularity assumptions $\boldsymbol{f}, \boldsymbol{u}_0 \in H_{\mathrm{div,per}}(\Omega)$ there exists a unique global attractor $\mathcal{A}_{\mathrm{Ler}}$ for the velocity. □

Remark 8.173 (Definition of the Viscous Dissipation Length Scale) The viscous dissipation length scale is the smallest scale that is needed to obtain a complete resolution of the flow field defined by the Leray-α model. The definition of this length scale is similar to the definition of the Kolmogorov length scale for the Navier–Stokes equations, see (8.3). Thus, one needs an expression for the dissipation of turbulent energy. It is remarked in Cheskidov et al. (2005) that a worst case scenario is given by

$$\varepsilon_{\mathrm{Ler}} = \frac{\nu}{(2\pi L)^3} \sup_{\boldsymbol{u}_0 \in \mathcal{A}_{\mathrm{Ler}}} \limsup_{t\to\infty} \frac{1}{t} \int_0^t \|\nabla \boldsymbol{w}(\tau)\|_{L^2(\Omega)}^2 \, d\tau.$$

Then, the viscous length scale of the Leray-α model is defined by

$$\lambda_{\nu,\mathrm{Ler}} = \left(\frac{\nu^3}{\varepsilon_{\mathrm{Ler}}}\right)^{1/4}.$$

Since the goal of the model consists in reducing the complexity of the Navier–Stokes equations, it can be expected that $\lambda_{\nu,\mathrm{Ler}} > \lambda$ or even $\lambda_{\nu,\mathrm{Ler}} \gg \lambda$. □

Remark 8.174 (The Dimension of the Attractor) To estimate the dimension of the global attractor, the Leray-α model is linearized about a trajectory in the attractor and the deviation is studied. The deviation satisfies a first order linear ordinary differential equation. There is a classical solution theory of this type of equations. In particular, it is well known that the norm of the solution (in an appropriate space) can be estimated by the norm of the initial condition times an exponential factor. In order that the deviation tends to zero, the argument of this exponential factor has to be negative. In Cheskidov et al. (2005), this argument is estimated and the condition that it should be negative gives an estimate of the Hausdorff dimension of the attractor

$$d_{\mathrm{H}}(\mathcal{A}_{\mathrm{Ler}}-) \le c \left(\frac{L}{\lambda_{\mathrm{Ler}}}\right)^{12/7} \left(1 + \frac{L}{\alpha}\right)^{9/14} \tag{8.194}$$

with some universal constant, which can be estimated explicitly. □

Remark 8.175 (Consequence of Estimate (8.194)) Already the fact that it is possible to prove the existence and uniqueness of a weak solution, Theorem 8.170, suggests that the Leray-α model is less complex than the Navier–Stokes equations. The estimate of the dimension of the attractor allows to quantify this suggestion to some extent.

The power $12/7$ of the attractor of the Leray-α model, for fixed α, is considerably smaller than the power 3 of the invariant bounded set X for the Navier–Stokes equations, see (8.10). In addition, one expects that $\lambda_{\text{Ler}} > \lambda$ or even $\lambda_{\text{Ler}} \gg \lambda$, see Remark 8.173. Altogether, (8.194) indicates that the number of degrees of freedom needed to simulate a flow modeled with the Leray-α model is much smaller than for a flow modeled with the Navier–Stokes equations.

In simulations, α will depend on the mesh width, see Remark 8.185. Inserting the estimate $\lambda_{\text{Ler}} \lesssim \alpha$ in (8.194) gives the estimate of the dimension $12/7 + 9/14 = 33/14 \in (2, 3)$. Also with this estimate, in combination with the expectation $\lambda_{\text{Ler}} > \lambda$ or $\lambda_{\text{Ler}} \gg \lambda$, the Leray-$\alpha$ model is an appropriate candidate for turbulence modeling. □

8.6.2 The Discrete Problem

Remark 8.176 (On the Application of the Differential Filter in Simulations) In numerical simulations, the differential filter as solution of a Helmholtz equation can be only approximated. For finite element methods, a straightforward approach consists in discretizing the Helmholtz equation in the velocity space, which gives the so-called discrete differential filter. The finite element error analysis requires some estimates of the discrete differential filter, which will be given next.

Since the differential filter requires to solve a Helmholtz equation for each component of the velocity separately, the analysis of the discrete differential filter will be presented for the scalar case. To avoid the introduction of further notations for this case, the continuous space will be denoted by $V = H_0^1(\Omega)$ and the conforming finite element space by $V^h \subset V$. □

Remark 8.177 (Differential Filter and Discrete Differential Filter) Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be a bounded domain with Lipschitz boundary Γ and $\alpha > 0$. Then the differential filter of $u \in V$ is the unique solution $\overline{u} \in V$ of

$$\alpha^2 (\nabla \overline{u}, \nabla v) + (\overline{u}, v) = (u, v) \quad \forall\ v \in V. \tag{8.195}$$

The discrete differential filter is the solution $\overline{u}^h \in V^h$ of

$$\alpha^2 \left(\nabla \overline{u}^h, \nabla v^h\right) + \left(\overline{u}^h, v^h\right) = \left(u, v^h\right) \quad \forall\ v^h \in V^h. \tag{8.196}$$

□

Lemma 8.178 (Properties of the Discrete Differential Filter) *The discrete differential filter is selfadjoint, i.e., it is*

$$\left(\overline{u^h}^h, v^h\right) = \left(u^h, \overline{v^h}^h\right) \quad \forall\, u^h, v^h \in V^h. \tag{8.197}$$

Let $u^h \in V^h$ be a time-dependent function, then it is $\partial_t \overline{u^h}^h = \overline{\partial_t u^h}^h \in V^h$, i.e., the filter and the differentiation in time commute.

Proof Since $\overline{u^h}^h \in V^h$, one can apply the definition (8.196) of the discrete differential filter with $\overline{u^h}^h$ as test function and obtains

$$\left(\overline{u^h}^h, v^h\right) = \left(\overline{u^h}^h, \overline{v^h}^h\right) + \alpha^2\left(\nabla \overline{u^h}^h, \nabla \overline{v^h}^h\right).$$

On the other hand, one gets with the test function $\overline{v^h}^h \in V^h$

$$\left(u^h, \overline{v^h}^h\right) = \left(\overline{u^h}^h, \overline{v^h}^h\right) + \alpha^2\left(\nabla \overline{u^h}^h, \nabla \overline{v^h}^h\right).$$

Combining these two equations proves (8.197).

Let $u^h \in V^h$ be given and let $v^h \in V^h$ be an arbitrary function, both functions might be time-dependent. Note that the filters of u^h, v^h, the temporal derivatives of u^h, v^h and their filters, and the filters of the temporal derivative are contained in V^h as well. Differentiating the definition (8.196) for u^h with respect to time, commuting integration in space and differentiation in time, applying the product rule, and commuting temporal and spatial derivatives gives

$$\begin{aligned}&\left(\partial_t u^h, v^h\right) + \left(u^h, \partial_t v^h\right)\\ &= \left(\partial_t \overline{u^h}^h, v^h\right) + \left(\overline{u^h}^h, \partial_t v^h\right) + \alpha^2\left(\nabla \partial_t \overline{u^h}^h, \nabla v^h\right) + \alpha^2\left(\nabla \overline{u^h}^h, \nabla \partial_t v^h\right).\end{aligned}$$

On the other hand, using the definition (8.196) of the discrete filter, one obtains

$$\begin{aligned}&\left(\partial_t u^h, v^h\right) + \left(u^h, \partial_t v^h\right)\\ &= \left(\overline{\partial_t u^h}^h, v^h\right) + \alpha^2\left(\nabla \overline{\partial_t u^h}^h, \nabla v^h\right) + \left(\overline{u^h}^h, \partial_t v^h\right) + \alpha^2\left(\nabla \overline{u^h}^h, \nabla \partial_t v^h\right).\end{aligned}$$

Combining these equations yields

$$\left(\partial_t \overline{u^h}^h - \overline{\partial_t u^h}^h, v^h\right) + \alpha^2\left(\nabla\left(\partial_t \overline{u^h}^h - \overline{\partial_t u^h}^h\right), \nabla v^h\right) = 0 \quad \forall\, v^h \in V^h.$$

Since the left-hand side of this equation defines an inner product in V^h, the assumptions of the Theorem of Lax–Milgram, Theorem B.4 are satisfied and applying this theorem it follows that $\partial_t \overline{u^h}^h - \overline{\partial_t u^h}^h = 0$ is the unique solution of this equation. ■

Lemma 8.179 (Stability of the Discrete Differential Filter) *Let $u \in V$, then it holds*

$$\left\| \overline{u}^h \right\|_{L^2(\Omega)} \le \|u\|_{L^2(\Omega)}, \tag{8.198}$$

$$\left\| \nabla \overline{u}^h \right\|_{L^2(\Omega)} \le \|\nabla u\|_{L^2(\Omega)}. \tag{8.199}$$

Proof Inserting $v^h = \overline{u}^h$ in (8.196), applying the Cauchy–Schwarz inequality (A.10), and Young's inequality (A.5) yields

$$\alpha^2 \left\| \nabla \overline{u}^h \right\|_{L^2(\Omega)}^2 + \left\| \overline{u}^h \right\|_{L^2(\Omega)}^2 \le \|u\|_{L^2(\Omega)} \left\| \overline{u}^h \right\|_{L^2(\Omega)} \le \frac{\|u\|_{L^2(\Omega)}^2}{2} + \frac{\left\| \overline{u}^h \right\|_{L^2(\Omega)}^2}{2}.$$

Bounding the first term on the left-hand side by zero from below gives immediately (8.198).

To prove the second estimate, one defines the discrete Laplacian $\Delta^h : V \to V^h$ by

$$\left(\Delta^h u, v^h\right) = -\left(\nabla u, \nabla v^h\right) \quad \forall\ v^h \in V^h. \tag{8.200}$$

Applying this definition to the first term of (8.196) gives

$$\left(\nabla \overline{u}^h, \nabla v^h\right) = -\left(\Delta^h \overline{u}^h, v^h\right).$$

Inserting now $v^h = \Delta^h \overline{u}^h$ in (8.196) leads to

$$-\alpha^2 \left\| \Delta^h \overline{u}^h \right\|_{L^2(\Omega)} + \left(\overline{u}^h, \Delta^h \overline{u}^h\right) = \left(u, \Delta^h \overline{u}^h\right),$$

such that with (8.200)

$$\alpha^2 \left\| \Delta^h \overline{u}^h \right\|_{L^2(\Omega)} + \left\| \nabla \overline{u}^h \right\|_{L^2(\Omega)} = \left(\nabla u, \nabla \overline{u}^h\right). \tag{8.201}$$

With the same reasoning as in the first part of the proof, one obtains now (8.199). ■

Lemma 8.180 (Error Estimate for the Discrete Differential Filter in Terms of u) *Let $u \in V$ with $\Delta u \in L^2(\Omega)$, then it is*

$$\begin{aligned} &\alpha^2 \left\| \nabla \left(u - \overline{u}^h\right) \right\|_{L^2(\Omega)}^2 + \left\| u - \overline{u}^h \right\|_{L^2(\Omega)}^2 \\ &\le C \left[\inf_{v^h \in V^h} \left(\alpha^2 \left\| \nabla \left(u - v^h\right) \right\|_{L^2(\Omega)}^2 + \left\| u - v^h \right\|_{L^2(\Omega)}^2 \right) + \alpha^4 \left\| \Delta u \right\|_{L^2(\Omega)}^2 \right], \end{aligned} \tag{8.202}$$

where C does not depend on α and h. Consequently, one has

$$\left\| \nabla \left(u - \overline{u}^h\right) \right\|_{L^2(\Omega)} \le \frac{C}{\alpha} \left((\alpha + h) \left\| \nabla u \right\|_{L^2(\Omega)} + \alpha^2 \left\| \Delta u \right\|_{L^2(\Omega)} \right), \qquad (8.203)$$

$$\left\| u - \overline{u}^h \right\|_{L^2(\Omega)} \le C \left((\alpha + h) \left\| \nabla u \right\|_{L^2(\Omega)} + \alpha^2 \left\| \Delta u \right\|_{L^2(\Omega)} \right). \qquad (8.204)$$

Proof The proof proceeds in principle along the standard lines of proving finite element error estimates. The only non-standard issue is that the continuous function in the error is also on the right-hand side of the discrete problem.

By the regularity assumption, u satisfies

$$\alpha^2 \left(\nabla u, \nabla v^h \right) + \left(u, v^h \right) = -\alpha^2 \left(\Delta u, v^h \right) + \left(u, v^h \right) \quad \forall\ v^h \in V^h.$$

Denoting the error by $e = u - \overline{u}^h$ and subtracting (8.196) from this identity yields an error equation

$$\alpha^2 \left(\nabla e, \nabla v^h \right) + \left(e, v^h \right) = -\alpha^2 \left(\Delta u, v^h \right) \quad \forall\ v^h \in V^h. \qquad (8.205)$$

Next, the error is decomposed into $e = \left(u - I^h u\right) - \left(\overline{u}^h - I^h u\right) = \eta - \varphi^h$, where $I^h u \in V^h$ is some arbitrary interpolant. Inserting this decomposition in the error equation and setting $v^h = \varphi^h$ gives

$$\alpha^2 \left\| \nabla \varphi^h \right\|_{L^2(\Omega)}^2 + \left\| \varphi^h \right\|_{L^2(\Omega)}^2 = \alpha^2 \left(\nabla \eta, \nabla \varphi^h \right) + \left(\eta, \varphi^h \right) + \alpha^2 \left(\Delta u, \varphi^h \right).$$

One obtains, applying the Cauchy–Schwarz inequality (A.10) and Young's inequality (A.5),

$$\begin{aligned}
&\alpha^2 \left\| \nabla \varphi^h \right\|_{L^2(\Omega)}^2 + \left\| \varphi^h \right\|_{L^2(\Omega)}^2 \\
&\le \alpha^2 \left\| \nabla \eta \right\|_{L^2(\Omega)} \left\| \nabla \varphi^h \right\|_{L^2(\Omega)} + \left\| \eta \right\|_{L^2(\Omega)} \left\| \varphi^h \right\|_{L^2(\Omega)} + \alpha^2 \left\| \Delta u \right\|_{L^2(\Omega)} \left\| \varphi^h \right\|_{L^2(\Omega)} \\
&\le \frac{\alpha^2}{2} \left\| \nabla \eta \right\|_{L^2(\Omega)}^2 + \frac{\alpha^2}{2} \left\| \nabla \varphi^h \right\|_{L^2(\Omega)}^2 + \left\| \eta \right\|_{L^2(\Omega)}^2 + \frac{1}{4} \left\| \varphi^h \right\|_{L^2(\Omega)}^2 \\
&\quad + \alpha^4 \left\| \Delta u \right\|_{L^2(\Omega)}^2 + \frac{1}{4} \left\| \varphi^h \right\|_{L^2(\Omega)}^2.
\end{aligned}$$

The terms with φ^h on the right-hand side can be absorbed from the left-hand side. Finally, an application of the triangle inequality gives

$$\begin{aligned}
&\alpha^2 \left\| \nabla e \right\|_{L^2(\Omega)}^2 + \left\| e \right\|_{L^2(\Omega)}^2 \\
&\le 2 \left(\alpha^2 \left\| \nabla \eta \right\|_{L^2(\Omega)}^2 + \left\| \eta \right\|_{L^2(\Omega)}^2 + \alpha^2 \left\| \nabla \varphi^h \right\|_{L^2(\Omega)}^2 + \left\| \varphi^h \right\|_{L^2(\Omega)}^2 \right) \\
&\le C \left(\alpha^2 \left\| \nabla \eta \right\|_{L^2(\Omega)}^2 + \left\| \eta \right\|_{L^2(\Omega)}^2 + \alpha^4 \left\| \Delta u \right\|_{L^2(\Omega)}^2 \right),
\end{aligned}$$

which completes the proof of (8.202), since $I^h u$ was chosen to be arbitrary.

The estimates (8.203) and (8.204) are derived by bounding the best approximation errors in (8.202) with interpolation errors and by applying the interpolation estimate (C.14). ■

Remark 8.181 (Finite Element Error Analysis of the Leray-α Model) The available finite element error analysis for the Leray-α model proceeds in the same way as for the Galerkin discretization of the Navier–Stokes equations, see Sect. 7.2. In particular, the obtained results are qualitatively not better than for the Galerkin discretization in the sense that the constants in the error bounds depend still on $\exp(C\nu^{-3})$ or on $\exp(C\nu^{-1})$, depending on the assumed regularity of the solution, where C depends on norms of the solution of the Navier–Stokes equations. For these reasons, a presentation of a detailed analysis would be just a repetition and it will be omitted here. Instead, only the differences in the finite element error analysis for the continuous-in-time case will be presented. The discussion of the obtained estimates gives a guideline for the asymptotic choice of the parameter α. □

Remark 8.182 (Continuous-in-Time Leray-α Finite Element Model) Consider inf-sup stable finite element spaces $V^h \subset V$ and $Q^h \subset Q$, then the continuous-in-time Leray-α finite element model reads as follows: Find $\boldsymbol{w}^h : (0,T] \to V^h$ and $r^h : (0,T] \to Q^h$ such that

$$\begin{aligned}\left(\partial_t \boldsymbol{w}^h, \boldsymbol{v}^h\right) + (\nu\nabla\boldsymbol{w}^h, \nabla\boldsymbol{v}^h) + n_{\text{skew}}\left(\overline{\boldsymbol{w}^h}^h, \boldsymbol{w}^h, \boldsymbol{v}^h\right)&\\ - \left(\nabla\cdot\boldsymbol{v}^h, r^h\right) + \left(\nabla\cdot\boldsymbol{w}^h, q^h\right) &= \langle \boldsymbol{f}, \boldsymbol{v}^h\rangle_{V',V},\\ \alpha^2\left(\nabla\overline{\boldsymbol{w}^h}^h, \nabla\boldsymbol{v}^h\right) + \left(\overline{\boldsymbol{w}^h}^h, \boldsymbol{v}^h\right) &= \left(\boldsymbol{w}^h, \boldsymbol{v}^h\right)\end{aligned} \tag{8.206}$$

for all $(\boldsymbol{v}^h, q^h) \in V^h \times Q^h$, $\alpha > 0$, and $\boldsymbol{w}^h(0,\boldsymbol{x}) \in V^h$ is an approximation of $\boldsymbol{u}_0(\boldsymbol{x})$. □

Lemma 8.183 (Existence, Uniqueness, and Stability of the Finite Element Solution) *Let $\boldsymbol{w}_0^h \in V_{\text{div}}^h$ and $\boldsymbol{f} \in L^2(0,t;V')$, then the finite element problem (8.206) has a unique solution $\left(\boldsymbol{w}^h, r^h\right) \in V^h \times Q^h$. For all $t \in (0,T]$, the stability estimate*

$$\left\|\boldsymbol{w}^h(t)\right\|_{L^2(\Omega)}^2 + \nu\left\|\nabla\boldsymbol{w}^h\right\|_{L^2(0,t;L^2(\Omega))}^2 \le \left\|\boldsymbol{w}_0^h\right\|_{L^2(\Omega)}^2 + \frac{1}{\nu}\left\|\boldsymbol{f}\right\|_{L^2(0,t;V')}^2 \tag{8.207}$$

holds.

Proof The existence and uniqueness of a solution can be proved in the same way as it is done in the first step of the Galerkin method for proving the existence of a weak solution, e.g., see Lemma 7.12.

To prove the stability estimate, choose $(\boldsymbol{v}^h, q^h) = (\boldsymbol{w}^h, r^h)$ in (8.206). Using (6.26), the nonlinear convective term vanishes. Now, the proof proceeds analogously as the proof of Lemma 7.12 starting from (7.14). Note that (8.207) is of the same form as estimate (7.11). In addition, an estimate of form (7.10) can be derived for $\boldsymbol{w}^h$ in the same way as in the proof of Lemma 7.12. ■

Theorem 8.184 (Finite Element Error Estimates for the Continuous-in-Time Leray-α Finite Element Model) *Let the assumptions of Theorem 7.35 be satisfied and let in addition*

$$\Delta \boldsymbol{u} \in L^4\left(0,T;L^2\left(\Omega\right)\right). \tag{8.208}$$

Then the following error estimate holds for all $t \in (0,T]$

$$\begin{aligned}
&\left\|\left(\boldsymbol{u}-\boldsymbol{w}^h\right)(t)\right\|_{L^2(\Omega)}^2 + \nu\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{w}^h\right)\right\|_{L^2(0,t;L^2(\Omega))}^2 \\
&\le \text{right-hand side of (7.44)} \\
&\quad + \frac{C}{\nu}\exp\left(\frac{C}{\nu^3}\left\|\nabla\boldsymbol{u}\right\|_{L^4(0,t;L^2(\Omega))}^4\right)\Bigg[\left(\alpha^{1/2}+\alpha^{-1/2}h\right)^2\left\|\nabla\boldsymbol{u}\right\|_{L^4(0,t;L^2(\Omega))}^4 \\
&\quad + \alpha^3\left\|\Delta\boldsymbol{u}\right\|_{L^4(0,t;L^2(\Omega))}^2\left\|\nabla\boldsymbol{u}\right\|_{L^4(0,t;L^2(\Omega))}^2\Bigg].
\end{aligned} \tag{8.209}$$

Assuming in addition to (8.208)

$$\boldsymbol{u} \in L^4\left(0,T;L^\infty\left(\Omega\right)\right), \quad \nabla\boldsymbol{u} \in L^4\left(0,T;L^\infty\left(\Omega\right)\right), \tag{8.210}$$

then one obtains the error estimate for all $t \in (0,T]$

$$\begin{aligned}
&\left\|\left(\boldsymbol{u}-\boldsymbol{w}^h\right)(t)\right\|_{L^2(\Omega)}^2 + \nu\left\|\nabla\left(\boldsymbol{u}-\boldsymbol{w}^h\right)\right\|_{L^2(0,t;L^2(\Omega))}^2 \\
&\le \text{similar to right-hand side of (7.44) with} \\
&\qquad\qquad \exp\left(C\left(\left\|\nabla\boldsymbol{u}\right\|_{L^\infty(\Omega)} + \frac{1}{\nu}\left\|\boldsymbol{u}\right\|_{L^\infty(\Omega)}^2\right)\right) \\
&\quad + \frac{C}{\nu}\exp\left(C\left(\left\|\nabla\boldsymbol{u}\right\|_{L^\infty(\Omega)} + \frac{1}{\nu}\left\|\boldsymbol{u}\right\|_{L^\infty(\Omega)}^2\right)\right)\Bigg[\left((\alpha+h)^2+\alpha^4\right) \\
&\quad \times\Big(\left\|\nabla\boldsymbol{u}\right\|_{L^4(0,t;L^2(\Omega))}^2 + \left\|\Delta\boldsymbol{u}\right\|_{L^4(0,t;L^2(\Omega))}^2 + \left\|\boldsymbol{u}\right\|_{L^4(0,t;L^\infty(\Omega))}^2 \\
&\quad + \left\|\nabla\boldsymbol{u}\right\|_{L^4(0,t;L^\infty(\Omega))}^2\Big)\Bigg].
\end{aligned} \tag{8.211}$$

Proof The proof follows exactly the lines of the proof of Theorem 7.35. As mentioned in Remark 8.181, only this part will be presented in detail which is different.

Using the same notations as in the proof of Theorem 7.35 and considering the decomposition

$$\boldsymbol{e}(t) = \boldsymbol{u}(t) - \boldsymbol{w}^h(t) = \left(\boldsymbol{u}(t) - I_{\mathrm{St}}^h\boldsymbol{u}(t)\right) + \left(I_{\mathrm{St}}^h\boldsymbol{u}(t) - \boldsymbol{w}^h(t)\right) = \boldsymbol{\eta}(t) - \boldsymbol{\phi}^h(t),$$

the critical term for the estimate is the difference of the nonlinear convective terms. One gets for the Leray-α model, using (6.26),

$$\begin{aligned}
& n_{\text{skew}}\left(\boldsymbol{u},\boldsymbol{u},\boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\overline{\boldsymbol{w}^h}^h,\boldsymbol{w}^h,\boldsymbol{\phi}^h\right) \\
&= n_{\text{skew}}\left(\boldsymbol{u}-\overline{\boldsymbol{u}}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right) + n_{\text{skew}}\left(\overline{\boldsymbol{u}}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\overline{\boldsymbol{w}^h}^h,\boldsymbol{w}^h,\boldsymbol{\phi}^h\right) \\
&= n_{\text{skew}}\left(\boldsymbol{u}-\overline{\boldsymbol{u}}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right) + n_{\text{skew}}\left(\overline{\boldsymbol{\eta}}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\overline{\boldsymbol{\phi}^h}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right) \\
&\quad + n_{\text{skew}}\left(\overline{\boldsymbol{w}^h}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\overline{\boldsymbol{w}^h}^h,\boldsymbol{w}^h,\boldsymbol{\phi}^h\right) \\
&= n_{\text{skew}}\left(\boldsymbol{u}-\overline{\boldsymbol{u}}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right) + n_{\text{skew}}\left(\overline{\boldsymbol{\eta}}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\overline{\boldsymbol{\phi}^h}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right) \\
&\quad + n_{\text{skew}}\left(\overline{\boldsymbol{w}^h}^h,\boldsymbol{u}-\boldsymbol{w}^h,\boldsymbol{\phi}^h\right) \\
&= n_{\text{skew}}\left(\boldsymbol{u}-\overline{\boldsymbol{u}}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right) + n_{\text{skew}}\left(\overline{\boldsymbol{\eta}}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right) - n_{\text{skew}}\left(\overline{\boldsymbol{\phi}^h}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right) \\
&\quad + n_{\text{skew}}\left(\overline{\boldsymbol{w}^h}^h,\boldsymbol{\eta},\boldsymbol{\phi}^h\right). \qquad (8.212)
\end{aligned}$$

The last three terms are estimated the same way as (7.50)–(7.52). Using the stability estimates (8.198) and (8.199) for the discrete filter, one gets even exactly the same estimates as in (7.50)–(7.52).

With the assumptions of Theorem 7.35, one can estimate the first term with (6.41) for $s = 1/2$, and Young's inequality (A.5)

$$\begin{aligned}
& n_{\text{skew}}\left(\boldsymbol{u}-\overline{\boldsymbol{u}}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right) \\
&\le C\left\|\boldsymbol{u}-\overline{\boldsymbol{u}}^h\right\|_{L^2(\Omega)}^{1/2}\left\|\nabla\left(\boldsymbol{u}-\overline{\boldsymbol{u}}^h\right)\right\|_{L^2(\Omega)}^{1/2}\left\|\nabla\boldsymbol{u}\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \\
&\le \frac{C}{\nu}\left\|\boldsymbol{u}-\overline{\boldsymbol{u}}^h\right\|_{L^2(\Omega)}\left\|\nabla\left(\boldsymbol{u}-\overline{\boldsymbol{u}}^h\right)\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{u}\right\|_{L^2(\Omega)}^2 + \frac{\nu}{16}\left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2.
\end{aligned}$$

The last term is absorbed in the left-hand side of the differential equation for the error estimate. The error estimate (8.204) gives for the first term an estimate of the form

$$\begin{aligned}
& \left\|\boldsymbol{u}-\overline{\boldsymbol{u}}^h\right\|_{L^2(\Omega)}\left\|\nabla\left(\boldsymbol{u}-\overline{\boldsymbol{u}}^h\right)\right\|_{L^2(\Omega)}\left\|\nabla\boldsymbol{u}\right\|_{L^2(\Omega)}^2 \\
&\le \frac{C}{\alpha}\left((\alpha+h)\left\|\nabla\boldsymbol{u}\right\|_{L^2(\Omega)} + \alpha^2\left\|\Delta\boldsymbol{u}\right\|_{L^2(\Omega)}\right)^2\left\|\nabla\boldsymbol{u}\right\|_{L^2(\Omega)}^2 \\
&\le \frac{C}{\alpha}(\alpha+h)^2\left\|\nabla\boldsymbol{u}\right\|_{L^2(\Omega)}^4 + C\alpha^3\left\|\Delta\boldsymbol{u}\right\|_{L^2(\Omega)}^2\left\|\nabla\boldsymbol{u}\right\|_{L^2(\Omega)}^2.
\end{aligned}$$

Using the Cauchy–Schwarz inequality (A.10), one has

$$\int_0^t \|\Delta \boldsymbol{u}\|_{L^2(\Omega)}^2 \|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^2 \, d\tau \\ \leq \left(\int_0^t \|\Delta \boldsymbol{u}\|_{L^2(\Omega)}^4 \, d\tau\right)^{1/2} \left(\int_0^t \|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^4 \, d\tau\right)^{1/2} < \infty, \tag{8.213}$$

by the regularity assumptions. Now, estimate (8.209) follows in the same way as in Theorem 7.35.

With assumption (8.210), one can estimate the first term of the last bound in (8.212) with (6.39), Poincaré's inequality (A.12), Young's inequality, and (8.204)

$$\begin{aligned} & n_{\text{skew}}\left(\boldsymbol{u} - \overline{\boldsymbol{u}}^h, \boldsymbol{u}, \boldsymbol{\phi}^h\right) \\ & \leq \frac{1}{2} \left\|\boldsymbol{u} - \overline{\boldsymbol{u}}^h\right\|_{L^2(\Omega)} \left(\|\nabla \boldsymbol{u}\|_{L^\infty(\Omega)} \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} + \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \|\boldsymbol{u}\|_{L^\infty(\Omega)}\right) \\ & \leq C \left\|\boldsymbol{u} - \overline{\boldsymbol{u}}^h\right\|_{L^2(\Omega)} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \left(\|\nabla \boldsymbol{u}\|_{L^\infty(\Omega)} + \|\boldsymbol{u}\|_{L^\infty(\Omega)}\right) \\ & \leq \frac{C}{\nu} \left\|\boldsymbol{u} - \overline{\boldsymbol{u}}^h\right\|_{L^2(\Omega)}^2 \left(\|\nabla \boldsymbol{u}\|_{L^\infty(\Omega)} + \|\boldsymbol{u}\|_{L^\infty(\Omega)}\right)^2 + \frac{\nu}{16} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 \\ & \leq \frac{C}{\nu} \left((\alpha + h)^2 \|\nabla \boldsymbol{u}\|_{L^2(\Omega)}^2 + \alpha^4 \|\Delta \boldsymbol{u}\|_{L^2(\Omega)}^2\right) \left(\|\nabla \boldsymbol{u}\|_{L^\infty(\Omega)} + \|\boldsymbol{u}\|_{L^\infty(\Omega)}\right)^2 \\ & \quad + \frac{\nu}{16} \left\|\nabla \boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 . \end{aligned} \tag{8.214}$$

The last term is absorbed from the left-hand side of the differential inequality. By assumptions (8.208) and (8.210), one shows analogously to (8.213) that all terms are in $L^1(0, T)$ such that Gronwall's lemma can be applied. With the assumed regularity of the solution, the third term on the right-hand side of (8.212) can be bounded in the form (7.58), using also the stability estimate (8.198). Thus, in this case, the term in the exponential depends only on ν^{-1}. Then, estimate (8.211) follows again in the same way as in Theorem 7.35. ■

Remark 8.185 (Optimal Asymptotic Choice of α) For both error bounds (8.209) and (8.211), $\alpha \sim h$ is the optimal asymptotic choice. In the case of (8.209), this choice follows from equilibrating the terms α and $\alpha^{-1}h^2$. One obtains first order convergence, also for higher order finite elements.

In (8.211), one can choose on the one hand $\alpha = 0$ to get the best error bound. But on the other hand, α should be as large as possible to have a sufficient impact of the turbulence model. The optimal compromise is $\alpha \sim h$. Then, one gets the same power for the terms α^2 and h^2 in the parentheses. This power cannot be improved by different choices. Altogether, there is a second order of convergence in this case. □

Remark 8.186 (The Fully Discrete Case) A finite element error analysis for the Leray-α finite element model discretized in time with the Crank–Nicolson scheme, see Remark 7.49, can be found in Layton et al. (2008). In fact, this analysis is the special case of $N = 0$ in this paper. The existence of a solution in each discrete time is proved and a stability estimate is derived. The finite element error analysis uses an estimate of type (8.214) for the nonlinear term which is introduced by the turbulence model. As in the continuous-in-time case, an error bound is derived where the constant depends on $\exp(C\nu^{-3})$, where C depends on norms of the solution of the Navier–Stokes equations. The application of the discrete Gronwall lemma, Lemma A.56, gives the usual time step restriction. Finally, error bounds are proved of second order in α, like in the corresponding situation for the continuous-in-time case, see Remark 8.185, and second order in Δt. □

Remark 8.187 (Leray-α Approximate Deconvolution Model) The idea of applying an approximate deconvolution, see Remark 8.150, can be also applied to the convective field of the Leray-α model, leading to the Leray-α ADM or Leray-deconvolution model

$$\begin{aligned} \partial_t \boldsymbol{w} - \nu\Delta\boldsymbol{w} + (G_{\mathrm{deconv},N}(\overline{\boldsymbol{w}})\cdot\nabla)\,\boldsymbol{w} + \nabla r &= \boldsymbol{f} \ \text{ in } (0,T]\times\Omega, \\ \nabla\cdot\boldsymbol{w} &= 0 \ \text{ in } (0,T]\times\Omega, \\ -\alpha^2\Delta\overline{\boldsymbol{w}} + \overline{\boldsymbol{w}} &= \boldsymbol{w} \ \text{in } (0,T]\times\Omega. \end{aligned}$$

This model was analyzed in the usual setup for ADMs, i.e., for space-periodic boundary conditions, the differential filter, and the van Cittert approximate deconvolution operator, see Layton and Rebholz (2012, Chap. 6) for details and an overview on available results. □

Example 8.188 (Turbulent Channel Flow at $Re_\tau = 180$) This benchmark problem is described in Example D.12. The setup of the simulations was almost identical to the setup for the Smagorinsky model, compare Example 8.128. The only difference was the initial condition. For the Leray-α model (8.206), with $n_{\mathrm{conv}}\left(\overline{\boldsymbol{w}^h}^h, \boldsymbol{w}^h, \boldsymbol{v}^h\right)$ instead of $n_{\mathrm{skew}}\left(\overline{\boldsymbol{w}^h}^h, \boldsymbol{w}^h, \boldsymbol{v}^h\right)$, a fully developed flow field was used. With the initial condition (D.34), it could be observed that the interval $[0, 10]$ was not sufficient for obtaining a fully developed flow field for some values of the model parameter.

Figure 8.9 presents the obtained results for different values of the model parameter $\alpha = \alpha_0 h_{K,\mathrm{short}}$. With respect to $\boldsymbol{u}_{\mathrm{mean}}$ there are only little differences for $\alpha_0 \in [0.4, 0.8]$. Considering in this interval the second order statistics of interest, then the best results were computed with $\alpha_0 = 0.8$. □

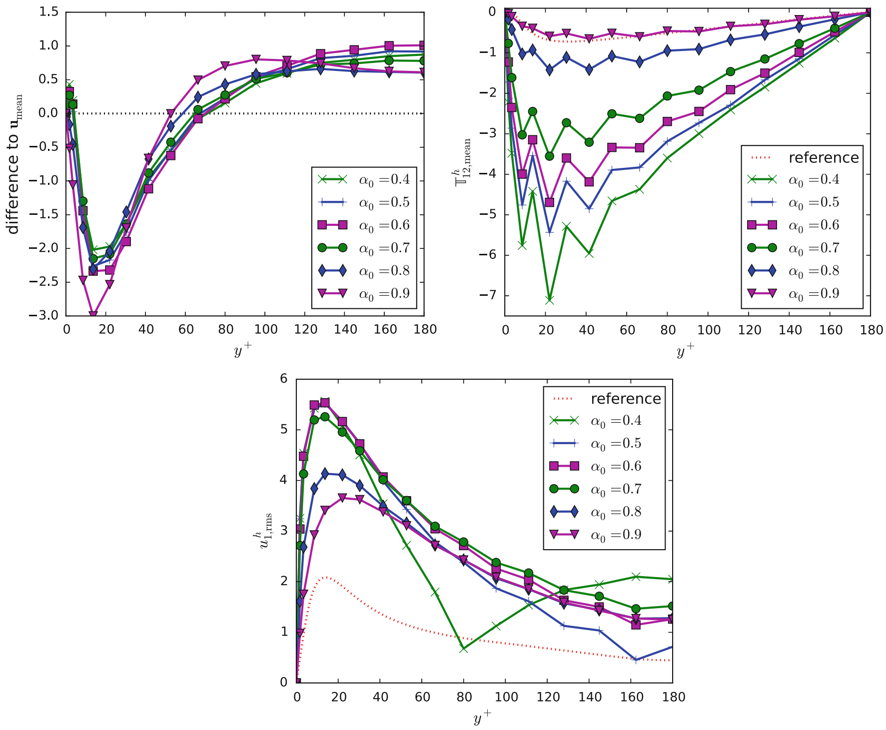

Fig. 8.9 Example 8.188. Statistics of interest for the Leray-α model with different values of $\alpha = \alpha_0 h_{K,\text{short}}$

8.7 The Navier–Stokes-α Model

Remark 8.189 (The Navier–Stokes-α Model) The Navier–Stokes-α model is given by

$$
\begin{aligned}
\partial_t \boldsymbol{w} - \nu \Delta \boldsymbol{w} + (\overline{\boldsymbol{w}} \cdot \nabla)\boldsymbol{w} + (\nabla \overline{\boldsymbol{w}})^T \boldsymbol{w} + \nabla r = \boldsymbol{f} \quad & \text{in } (0,T] \times \Omega, \\
\nabla \cdot \overline{\boldsymbol{w}} = 0 \quad & \text{in } (0,T] \times \Omega, \\
\boldsymbol{w}(0,\cdot) = \boldsymbol{u}_0 \quad & \text{in } \Omega, \\
-\alpha^2 \Delta \overline{\boldsymbol{w}} + \overline{\boldsymbol{w}} = \boldsymbol{w} \quad & \text{in } (0,T] \times \Omega, \\
\int_\Omega r \, d\boldsymbol{x} = 0 \quad & \text{in } (0,T],
\end{aligned}
\tag{8.215}
$$

together with appropriate boundary conditions. The pressure includes some terms that appear in the derivation of this model and it has the form, Chen et al. (1998),

$$
r = p - \frac{1}{2} \left\| \overline{\boldsymbol{w}} \right\|_2^2 - \frac{1}{2}\alpha^2 \left(\nabla \overline{\boldsymbol{w}} : \overline{\boldsymbol{w}} \right).
$$

Model (8.215) is also called viscous Camassa–Holm model or isotropic Lagrangian-averaged Navier–Stokes equations.

Using (3.158), the momentum equation can be rewritten, introducing a new pressure, by

$$\partial_t \boldsymbol{w} - \nu \Delta \boldsymbol{w} + (\nabla \times \boldsymbol{w}) \times \overline{\boldsymbol{w}} + \nabla \tilde{r} = \boldsymbol{f} \tag{8.216}$$

or

$$\partial_t \boldsymbol{w} - \nu \Delta \boldsymbol{w} - \overline{\boldsymbol{w}} \times (\nabla \times \boldsymbol{w}) + \nabla \tilde{r} = \boldsymbol{f}. \tag{8.217}$$

□

Remark 8.190 (The Lagrangian Description of a Flow Field) The description of the flow field in Chap. 2, which finally led to the Navier–Stokes equations, was performed from the point of view of considering a point $(t, \boldsymbol{x})$ in time and space. Then, the flow field was modeled with functions depending on $(t, \boldsymbol{x})$. This approach is called the Eulerian description of the flow field. Alternatively, it is possible to describe the flow from the point of view of a "fluid particle". In this case, one follows the motion of that fluid particle through time and space. This approach is called Lagrangian description of a flow field.

The Eulerian description can be thought of sitting at the banks of a river and describing the flow of the river from this point. For the Lagrangian description, one sits in a boat and describes the flow by following the river with the boat. □

Remark 8.191 (Sketch of the Derivation) Although the Navier–Stokes-α model has a convective term with regularized velocity, its derivation is not based on regularization. One can find in the literature two ways for deriving (8.215).

One way, whose details can be found in Chen et al. (1999b), considers a Lagrangian functional comprised of the kinetic energy and the incompressibility constraint

$$\int \left[\frac{1}{2} \|\boldsymbol{u}(t, \boldsymbol{x})\|_2^2 + p\left(\boldsymbol{X}(t, \boldsymbol{x}), t\right) \left(\det\left(\nabla \boldsymbol{X}\left(t, \boldsymbol{x}\right)\right) - 1\right) \right] d\boldsymbol{x}$$

with $\boldsymbol{u}(t, \boldsymbol{x}) = \partial_t \boldsymbol{X}(t, \boldsymbol{x})$ and $\boldsymbol{X}(t, \boldsymbol{x})$ is the Lagrangian trajectory. The incompressibility constraint leads finally to the requirement $\det\left(\nabla \boldsymbol{X}\left(t, \boldsymbol{x}\right)\right) = 1$. Then, the Lagrangian trajectory is augmented with fluctuations. This step resembles the decomposition (8.11) of the flow field: the Lagrangian trajectory represents the mean flow field and the fluctuations the small scales. The trajectory with fluctuations is inserted in the Lagrangian functional. As a next crucial step, the velocity field and the pressure in this functional are approximated with a linear Taylor series expansion. This step assumes that the fluctuations are sufficiently small. Then, the Lagrangian functional is averaged and minimized. The optimality conditions are derived by computing its variational derivatives. In this way, one obtains a similar model to (8.215), but without viscous term and with a more complicated relation between $\overline{w}$ and w. Adding the viscous term, to be able to use this equation as a turbulence model for incompressible flow problems, was proposed in Chen et al. (1998, 1999a,b). If the fluctuations are assumed to be isotropic, one obtains the equation for $\overline{w}$ given in (8.215). If they are in addition homogeneous, then α is a constant. This derivation generalizes a one-dimensional shallow water model from Camassa and Holm (1993) to d dimensions. Therefore, (8.215) is called also viscous Camassa–Holm model.

A different derivation of the model (8.215) can be found in Marsden and Shkoller (2001, 2003). There, a so-called Lagrangian average is applied over a set of solutions of the Euler equations with initial data in some ball. The viscous term is treated via stochastic variations. It is noted that in the case of a bounded domain the definition of the viscous term should include a projection onto divergence-free vector fields. □

Remark 8.192 (The Divergence Constraint) If w and $\overline{w}$ are sufficiently smooth, one obtains from the definition of $\overline{w}$ in (8.215), $\nabla \cdot \overline{w} = 0$, and a calculation like in (8.187) gives

$$\nabla \cdot w = \nabla \cdot \overline{w} - \alpha^2 \nabla \cdot \Delta \overline{w} = \nabla \cdot \overline{w} - \alpha^2 \Delta (\nabla \cdot \overline{w}) = 0.$$

Hence, w is also divergence-free. □

Remark 8.193 (Analysis of the Navier–Stokes-α Model in Turbulent Channel and Pipe Flows) The Navier–Stokes-α model for turbulent channel and pipe flows was studied analytically in Chen et al. (1998, 1999a,b). It was found that the analytical steady-state solution of the Navier–Stokes-α model with constant α shows a good agreement, e.g., with available mean velocity profiles away from the boundary of a distance of order α. Note that the Navier–Stokes-α model was derived with the assumption of homogeneous and isotropic fluctuations, see Remark 8.191, which is usually not satisfied close to the boundary. Scaling arguments suggest that near the boundary α should decrease as the Reynolds number increases. The decrease of α, and with that the smaller influence of the turbulence model, resembles the reduction of the eddy viscosity in the Smagorinsky model by applying the van Driest damping, see Remark 8.127. Away from the boundary, the scaling arguments imply that α is independent of the Reynolds number. □

Remark 8.194 (Well-Posedness of the Navier–Stokes-α Model) One can find results on the existence and uniqueness of a solution of the Navier–Stokes-α model in Foias et al. (2002) for the case of periodic boundary conditions and in Marsden and Shkoller (2001) for the case of a bounded domain with no-slip boundary conditions. In Marsden and Shkoller (2001), first local well-posedness is proved by using Banach's fixed point theorem, see Remark 7.26 for the notation of a local solution. Then, global well-posedness is obtained by proving appropriate stability (a priori) estimates. The analysis in Foias et al. (2002) uses the Galerkin method presented in Sect. 7.1. It will be sketched here. □

Remark 8.195 (The Navier–Stokes-α Model in a Periodic Domain) Let $\Omega = (0,L)^3$, then the Navier–Stokes-α model with rotational form of the convective term has the form, see (8.217),

$$\begin{aligned} \partial_t \boldsymbol{w} - \nu\Delta\boldsymbol{w} - \overline{\boldsymbol{w}} \times (\nabla\times\boldsymbol{w}) + \nabla r &= \boldsymbol{f} && \text{in } (0,T]\times\Omega,\\ (\boldsymbol{w}, r) \text{ is periodic} & && \text{on } (0,T]\times\Gamma,\\ -\alpha^2\Delta\overline{\boldsymbol{w}} + \overline{\boldsymbol{w}} &= \boldsymbol{w} && \text{in } (0,T]\times\Omega,\\ \nabla\cdot\overline{\boldsymbol{w}} &= 0 && \text{in } (0,T]\times\Omega,\\ \overline{\boldsymbol{w}}(0,\cdot) &= \overline{\boldsymbol{u}_0} && \text{in } \Omega,\\ \overline{\boldsymbol{w}} \text{ is periodic} & && \text{on } (0,T]\times\Gamma,\\ \int_\Omega r\, d\boldsymbol{x} &= 0 && \text{in } (0,T], \end{aligned} \tag{8.218}$$

where for simplicity of notation the tilde is omitted for the pressure. Note that it does not matter if the initial condition for $\boldsymbol{w}$ or $\overline{\boldsymbol{w}}$ is prescribed. If one is known, the other one can be computed. □

Remark 8.196 (Assumptions and Consequences) In the analysis, it will be assumed that the right-hand side does not depend on time, i.e., $\boldsymbol{f}(t,\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})$. Further, it will be assumed that the right-hand side and the initial condition have zero mean, i.e.,

$$\int_\Omega \overline{\boldsymbol{w}}(0,\boldsymbol{x})\, d\boldsymbol{x} = \int_\Omega \boldsymbol{f}(\boldsymbol{x})\, d\boldsymbol{x} = 0. \tag{8.219}$$

Integration by parts gives

$$\begin{aligned} \int_\Omega \nabla r\, d\boldsymbol{x} &= -\int_{\partial\Omega} r\boldsymbol{n}\, d\boldsymbol{s}\\ &= \int_{x=0\cap\partial\Omega} r(-\boldsymbol{e}_1)\, d\boldsymbol{s} + \int_{x=L\cap\partial\Omega} r\boldsymbol{e}_1\, d\boldsymbol{s} + \int_{y=0\cap\partial\Omega} r(-\boldsymbol{e}_2)\, d\boldsymbol{s}\\ &\quad + \int_{y=L\cap\partial\Omega} r\boldsymbol{e}_2\, d\boldsymbol{s} + \int_{z=0\cap\partial\Omega} r(-\boldsymbol{e}_3)\, d\boldsymbol{s} + \int_{z=L\cap\partial\Omega} r\boldsymbol{e}_3\, d\boldsymbol{s}\\ &= 0, \end{aligned}$$

since r is periodic. With the same argument, one can derive

$$\int_\Omega \Delta\boldsymbol{w}\, d\boldsymbol{x} = \int_\Omega \overline{\boldsymbol{w}} \times (\nabla\times\boldsymbol{w})\, d\boldsymbol{x} = \int_\Omega \Delta\overline{\boldsymbol{w}}\, d\boldsymbol{x} = 0,$$

such that one finds from (8.218) and (8.219)

$$\begin{aligned} 0 &= \int_\Omega \boldsymbol{f}\, d\boldsymbol{x} = \int_\Omega \partial_t\boldsymbol{w}\, d\boldsymbol{x} = \frac{d}{dt}\int_\Omega \boldsymbol{w}\, d\boldsymbol{x}\\ &= \frac{d}{dt}\left(-\alpha^2\int_\Omega \Delta\overline{\boldsymbol{w}}\, d\boldsymbol{x} + \int_\Omega \overline{\boldsymbol{w}}\, d\boldsymbol{x}\right) = \frac{d}{dt}\int_\Omega \overline{\boldsymbol{w}}\, d\boldsymbol{x}. \end{aligned}$$

Thus, the last integral is a constant with respect to time and since this constant is zero at the initial time, see (8.219), it follows that

$$\int_{\Omega} \overline{\boldsymbol{w}}\ d\boldsymbol{x} = 0 \quad \forall\ t \geq 0.$$

□

Remark 8.197 (Setup for the Analysis) Let

$$\mathcal{V} = \left\{ \boldsymbol{v}\ :\ \boldsymbol{v} \text{ is a trigonometric polynomial on } \Omega \text{ with } \nabla \cdot \boldsymbol{v} = 0, \int_{\Omega} \boldsymbol{v}\ d\boldsymbol{x} = 0 \right\}.$$

In comparison with the situation in the non-periodic case, the condition with the vanishing mean value appears. Nevertheless, the same notations will be used here, namely $H_{\text{div}}(\Omega)$ is the completion of $\mathcal{V}$ in $L^2(\Omega)$ and V_{div} is the completion in $H^1(\Omega)$.

The Helmholtz projector $P_{\text{helm}}\ :\ L_0^2(\Omega) \to H_{\text{div}}(\Omega)$ is given in Definition 3.170. In addition, $A = -P_{\text{helm}}\Delta$ is the Stokes operator with domain $D(A) = H^2(\Omega) \cap V_{\text{div}}$.

Now, (8.218) can be written in operator form in the divergence-free subspace as follows

$$\begin{aligned} \frac{d}{dt}\boldsymbol{w} + \nu A\boldsymbol{w} + N(\overline{\boldsymbol{w}}, \boldsymbol{w}) &= P_{\text{helm}}\boldsymbol{f}, \\ \alpha^2 A\overline{\boldsymbol{w}} + \overline{\boldsymbol{w}} &= \boldsymbol{w}, \\ \overline{\boldsymbol{w}}(0) &= \overline{\boldsymbol{u}_0}, \end{aligned} \tag{8.220}$$

with $N(\overline{\boldsymbol{w}}, \boldsymbol{w}) = -P_{\text{helm}}(\overline{\boldsymbol{w}} \times (\nabla \times \boldsymbol{w}))$. □

Remark 8.198 (On the Stokes Operator) In the case of periodic boundary conditions, the restriction of A to $D(A)$ is a selfadjoint operator with compact inverse, e.g., see Temam (1995, p. 10). The Stokes operator possesses a set of eigenfunctions, which form an orthonormal basis of $H_{\text{div}}(\Omega)$, with corresponding positive eigenvalues, compare (Temam 1995, p. 10). □

Lemma 8.199 (Norm Estimates) *Let λ_1 be the smallest eigenvalue of the Stokes operator, then there holds the Poincaré-type inequality*

$$\|\boldsymbol{v}\|_{L^2(\Omega)}^2 \leq \lambda_1^{-1} \|\nabla \boldsymbol{v}\|_{L^2(\Omega)}^2 \quad \forall\ \boldsymbol{v} \in D(A). \tag{8.221}$$

For the Stokes operator, the following norm equivalence is valid

$$\|A\boldsymbol{v}\|_{L^2(\Omega)} \leq \|\boldsymbol{v}\|_{H^2(\Omega)} \leq C\|A\boldsymbol{v}\|_{L^2(\Omega)} \quad \forall\ \boldsymbol{v} \in D(A). \tag{8.222}$$

Proof Let $\{\boldsymbol{v}_l\}_{l=1}^{\infty}$ be the basis of orthonormal eigenfunctions of the Stokes operator and let $\boldsymbol{v} = \sum_{i=l}^{\infty} v_l \boldsymbol{v}_l \in D(A)$. Using an argument as in (8.187) shows that $\Delta \boldsymbol{v}$

is divergence-free, hence $-\Delta \boldsymbol{v} = A\boldsymbol{v}$. Then, one gets with integration by parts, the argument from the previous sentence, the definition of the eigenvalues, the orthonormality of the eigenfunctions, the positivity of the eigenvalues, and once more the orthonormality of the eigenfunctions

$$\begin{aligned}\|\nabla \boldsymbol{v}\|^2_{L^2(\Omega)} &= (\nabla \boldsymbol{v}, \nabla \boldsymbol{v}) = -(\boldsymbol{v}, \Delta \boldsymbol{v}) = (\boldsymbol{v}, A\boldsymbol{v}) \\ &= \left(\sum_{i=l}^{\infty} v_l \boldsymbol{v}_l, A \sum_{i=l}^{\infty} v_l \boldsymbol{v}_l\right) = \left(\sum_{i=l}^{\infty} v_l \boldsymbol{v}_l, \sum_{i=l}^{\infty} v_l \lambda_l \boldsymbol{v}_l\right) = \int_\Omega \sum_{i=l}^{\infty} v_l^2 \lambda_l \boldsymbol{v}_l \cdot \boldsymbol{v}_l \, d\boldsymbol{x} \\ &\geq \lambda_1 \int_\Omega \sum_{i=l}^{\infty} v_l^2 \boldsymbol{v}_l \cdot \boldsymbol{v}_l \, d\boldsymbol{x} = \lambda_1 \left(\sum_{i=l}^{\infty} v_l \boldsymbol{v}_l, \sum_{i=l}^{\infty} v_l \boldsymbol{v}_l\right) = \lambda_1 \|\boldsymbol{v}\|^2_{L^2(\Omega)} .\end{aligned}$$

For the proof of the norm equivalence (8.222), it is referred to Temam (1995, p. 9). The proof is based on the observation that the operator A defines an isomorphism between $D(A)$ and the space of all functions from $H^2(\Omega)$ whose mean value and the mean value of the derivatives vanish. ■

Lemma 8.200 (Estimate of the Convective Term) *Let* $\boldsymbol{u} \in V_{\text{div}}$, $\boldsymbol{v} \in H_{\text{div}}(\Omega)$, *and* $\boldsymbol{w} \in D(A)$, *then it is*

$$\begin{aligned}&|(-P_{\text{helm}} (\boldsymbol{u} \times (\nabla \times \boldsymbol{v})), \boldsymbol{w})| \\ &\leq C\Big(\|\boldsymbol{w}\|^{1/2}_{L^2(\Omega)} \|\nabla \boldsymbol{w}\|^{1/2}_{L^2(\Omega)} \|A\boldsymbol{u}\|_{L^2(\Omega)} \|\boldsymbol{v}\|_{L^2(\Omega)} \\ &\quad + \|\nabla \boldsymbol{u}\|^{1/2}_{L^2(\Omega)} \|A\boldsymbol{u}\|^{1/2}_{L^2(\Omega)} \|\nabla \boldsymbol{w}\|_{L^2(\Omega)} \|\boldsymbol{v}\|_{L^2(\Omega)} \Big). \qquad (8.223)\end{aligned}$$

Proof Since the Helmholtz projection is just the $L^2(\Omega)$ projection into $H_{\text{div}}(\Omega)$ and in particular $\boldsymbol{w} \in H_{\text{div}}(\Omega)$ by the definition of $D(A)$, one gets

$$(-P_{\text{helm}} (\boldsymbol{u} \times (\nabla \times \boldsymbol{v})), \boldsymbol{w}) = (-\boldsymbol{u} \times (\nabla \times \boldsymbol{v}), \boldsymbol{w}) = ((\nabla \times \boldsymbol{v}) \times \boldsymbol{u}, \boldsymbol{w}) .$$

Equality (6.14) can be derived also in the case of periodic boundary conditions. The integrals on the boundary, which appear in this derivation, will vanish because all functions are periodic and thus the integrals on opposite faces sum up to zero. Hence, one gets

$$(-P_{\text{helm}} (\boldsymbol{u} \times (\nabla \times \boldsymbol{w})), \boldsymbol{u}) = n_{\text{conv}} (\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) - n_{\text{conv}} (\boldsymbol{w}, \boldsymbol{v}, \boldsymbol{u}) . \qquad (8.224)$$

Also in the periodic case, the convective term is skew-symmetric, i.e., (6.24) holds. This property is derived as in Remark 6.8, where the boundary integral vanishes again because of the periodicity of the functions. With the triangle inequality, the skew-symmetry of the convective term, and Hölder's inequality (A.9), one gets

$$\begin{aligned}
&|(-P_{\text{helm}}(\boldsymbol{u}\times(\nabla\times\boldsymbol{v})),\boldsymbol{w})| \\
&\le |n_{\text{conv}}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})| + |n_{\text{conv}}(\boldsymbol{w},\boldsymbol{v},\boldsymbol{u})| \\
&\le |n_{\text{conv}}(\boldsymbol{u},\boldsymbol{w},\boldsymbol{v})| + |n_{\text{conv}}(\boldsymbol{w},\boldsymbol{u},\boldsymbol{v})| \\
&\le \|\boldsymbol{u}\|_{L^\infty(\Omega)}\|\nabla\boldsymbol{w}\|_{L^2(\Omega)}\|\boldsymbol{v}\|_{L^2(\Omega)} + \|\boldsymbol{w}\|_{L^3(\Omega)}\|\nabla\boldsymbol{u}\|_{L^6(\Omega)}\|\boldsymbol{v}\|_{L^2(\Omega)}. \quad (8.225)
\end{aligned}$$

Using the Sobolev imbedding (A.15), the interpolation theorem in Sobolev spaces (A.13), and Poincaré's inequality (A.12) gives

$$\|\boldsymbol{w}\|_{L^3(\Omega)} \le C\|\boldsymbol{w}\|_{H^{1/2}(\Omega)} \le C\|\boldsymbol{w}\|_{L^2(\Omega)}^{1/2}\|\nabla\boldsymbol{w}\|_{L^2(\Omega)}^{1/2}.$$

From the Sobolev imbedding (A.14) with $m = 1, p = 2, j = 1$, and with the norm equivalence (8.222), it follows that

$$\|\nabla\boldsymbol{u}\|_{L^6(\Omega)} \le C\|\boldsymbol{u}\|_{H^2(\Omega)} \le C\|A\boldsymbol{u}\|_{L^2(\Omega)}.$$

Using Agmon's inequality, see Foias et al. (2001, (A.29)), Poincaré's inequality, and the equivalence (8.222) yields

$$\|\boldsymbol{u}\|_{L^\infty(\Omega)} \le C\|\boldsymbol{u}\|_{H^1(\Omega)}^{1/2}\|\boldsymbol{u}\|_{H^2(\Omega)}^{1/2} \le C\|\nabla\boldsymbol{u}\|_{L^2(\Omega)}^{1/2}\|A\boldsymbol{u}\|_{L^2(\Omega)}^{1/2}.$$

Inserting all estimates in (8.225) gives the statement of the lemma. ■

Theorem 8.201 (Existence and Uniqueness of a Solution in the Space-Periodic Case) *Let* $\boldsymbol{f} \in H_{\text{div}}(\Omega)$ *and* $\overline{\boldsymbol{u}_0} \in V_{\text{div}}$. *Then there is for any* $T > 0$ *a unique solution of* (8.220)

$$\boldsymbol{w} \in L^\infty_{\text{loc}}\left((0,T], H^3(\Omega)\right).$$

Proof The proof utilizes the Galerkin method presented in Sect. 7.1. It consists of four parts:

1. Show the well-posedness of the problem in a finite-dimensional subspace.
2. Prove stability estimates in $H^1(\Omega), H^2(\Omega)$, and $H^3(\Omega)$.
3. Pass to the limit with the dimension and show the convergence of a subsequence.
4. Prove uniqueness of the solution.

The first three steps will be only sketched here.

1. *Show the well-posedness of the problem in a finite-dimensional subspace.* Analogously to the proof of Lemma 7.12, a problem in a finite-dimensional space is considered. The space is spanned by the eigenfunctions of the Stokes operator. The existence and uniqueness of an absolutely continuous solution in $[0,T]$ is proved with the theorem of Carathéodory, see Theorem A.50.

2. *Prove stability estimates in $H^1(\Omega), H^2(\Omega)$, and $H^3(\Omega)$.* The derivation of these estimates is somewhat longer. It uses standard estimates like Hölder's and Young's inequality, the Gronwall lemma, and estimates for the Stokes operator A and the operator N for the nonlinear term. Let $\overline{\boldsymbol{w}}^{\,n}(t)$ be the solution of the problem in the subspace with dimension n. Then, one obtains for all $t \in [0, T]$, e.g.,

$$\|\overline{\boldsymbol{w}}^{\,n}(t)\|_{L_2(\Omega)}^2 + \alpha^2 \|\nabla \overline{\boldsymbol{w}}^{\,n}(t)\|_{L_2(\Omega)}^2 \le C_1,$$
$$\|\nabla \overline{\boldsymbol{w}}^{\,n}(t)\|_{L_2(\Omega)}^2 + \alpha^2 \|A \overline{\boldsymbol{w}}^{\,n}(t)\|_{L_2(\Omega)}^2 \le C_2(t).$$

Note that there is the norm equivalence of $\|A \overline{\boldsymbol{w}}^{\,n}(t)\|_{L_2(\Omega)}$ and the $H^2(\Omega)$ norm, see (8.222).
3. *Pass to the limit with the dimension and show the convergence of a subsequences.* This part of the proof is performed like the proofs of Corollary 7.13 and Lemma 7.16, using the stability estimates from the previous part and the theorem of Lions–Aubin, Girault and Raviart (1979, p. 153)). Because of the higher regularity of the solution proved in the second step, in comparison with the solution of the Navier–Stokes equations, the proof of the convergence of the nonlinear convective term is simpler than in Sect. 7.1.
4. *Prove uniqueness of the solution.* Assume that there are two solutions $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ of (8.220) to the same data $\boldsymbol{f}$ and $\boldsymbol{u}_0$. Denoting $\boldsymbol{w}_{21} = \boldsymbol{w}_2 - \boldsymbol{w}_1$ and correspondingly $\overline{\boldsymbol{w}_{21}} = \overline{\boldsymbol{w}_2} - \overline{\boldsymbol{w}_1}$, one obtains by subtracting (8.220) for $\boldsymbol{w}_1$ from (8.220) for $\boldsymbol{w}_2$ and expanding with $N(\overline{\boldsymbol{w}_1}, \boldsymbol{w}_2) - N(\overline{\boldsymbol{w}_1}, \boldsymbol{w}_2)$

$$\begin{aligned} 0 &= \frac{d}{dt}\boldsymbol{w}_{21} + \nu A\boldsymbol{w}_{21} + N(\overline{\boldsymbol{w}_2}, \boldsymbol{w}_2) - N(\overline{\boldsymbol{w}_1}, \boldsymbol{w}_1) \\ &= \frac{d}{dt}\boldsymbol{w}_{21} + \nu A\boldsymbol{w}_{21} + N(\overline{\boldsymbol{w}_{21}}, \boldsymbol{w}_2) - N(\overline{\boldsymbol{w}_1}, \boldsymbol{w}_{21}). \end{aligned} \tag{8.226}$$

The next step consists in testing (8.226) with $\overline{\boldsymbol{w}_{21}}$. One gets for the first term on the right-hand side with the definition of the $\overline{\boldsymbol{w}_{21}}$, integration by parts, and relations of the form (7.13)

$$\begin{aligned} \left(\frac{d}{dt}\boldsymbol{w}_{21}, \overline{\boldsymbol{w}_{21}}\right) &= \frac{d}{dt}\left(\alpha^2 A \overline{\boldsymbol{w}_{21}} + \overline{\boldsymbol{w}_{21}}, \overline{\boldsymbol{w}_{21}}\right) \\ &= \alpha^2 \frac{d}{dt}\left(\nabla \overline{\boldsymbol{w}_{21}}, \nabla \overline{\boldsymbol{w}_{21}}\right) + \frac{d}{dt}\left(\overline{\boldsymbol{w}_{21}}, \overline{\boldsymbol{w}_{21}}\right) \\ &= \frac{1}{2}\frac{d}{dt}\left(\|\overline{\boldsymbol{w}_{21}}\|_{L^2(\Omega)}^2 + \alpha^2 \left\|\nabla \overline{\boldsymbol{w}_{21}}\right\|_{L^2(\Omega)}^2\right). \end{aligned}$$

For the viscous term, one obtains with applying repeatedly integration by parts

$$
\begin{aligned}
(\nu A \boldsymbol{w}_{21}, \overline{\boldsymbol{w}_{21}}) &= \alpha^2 \left(\nu A \left(A \overline{\boldsymbol{w}_{21}}\right), \overline{\boldsymbol{w}_{21}}\right) + \left(\nu A \overline{\boldsymbol{w}_{21}}, \overline{\boldsymbol{w}_{21}}\right) \\
&= \alpha^2 \left(\nu A \overline{\boldsymbol{w}_{21}}, A \overline{\boldsymbol{w}_{21}}\right) + \left(\nu \nabla \overline{\boldsymbol{w}_{21}}, \nabla \overline{\boldsymbol{w}_{21}}\right) \\
&= \nu \left(\alpha^2 \|A \overline{\boldsymbol{w}_{21}}\|_{L^2(\Omega)}^2 + \left\| \nabla \overline{\boldsymbol{w}_{21}} \right\|_{L^2(\Omega)}^2 \right).
\end{aligned}
$$

One finds with (8.224) that

$$
\begin{aligned}
\left(N\left(\overline{\boldsymbol{w}_{21}}, \boldsymbol{w}_2\right), \overline{\boldsymbol{w}_{21}}\right) &= \left(-P_{\text{helm}}\left(\overline{\boldsymbol{w}_{21}} \times (\nabla \times \boldsymbol{w}_2)\right), \overline{\boldsymbol{w}_{21}}\right) \\
&= n_{\text{conv}}\left(\overline{\boldsymbol{w}_{21}}, \boldsymbol{w}_2, \overline{\boldsymbol{w}_{21}}\right) - n_{\text{conv}}\left(\overline{\boldsymbol{w}_{21}}, \boldsymbol{w}_2, \overline{\boldsymbol{w}_{21}}\right) = 0.
\end{aligned}
$$

Inserting all identities in (8.226), applying estimate (8.223), collecting all norms of $\boldsymbol{w}_1$, which are known to be finite, into the constant, using (8.221), and Young's inequality (A.5) gives

$$
\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\left(\|\overline{\boldsymbol{w}_{21}}\|_{L^2(\Omega)}^2 + \alpha^2 \left\|\nabla \overline{\boldsymbol{w}_{21}}\right\|_{L^2(\Omega)}^2\right) + \nu\left(\alpha^2 \|A \overline{\boldsymbol{w}_{21}}\|_{L^2(\Omega)}^2 + \left\|\nabla \overline{\boldsymbol{w}_{21}}\right\|_{L^2(\Omega)}^2\right) \\
&\quad \le \left|\left(-N\left(\overline{\boldsymbol{w}_1}, \boldsymbol{w}_{21}\right), \overline{\boldsymbol{w}_{21}}\right)\right| \\
&\quad = \left|\left(-P_{\text{helm}}\left(\boldsymbol{w}_1 \times (\nabla \times \boldsymbol{w}_{21})\right), \overline{\boldsymbol{w}_{21}}\right)\right| \\
&\quad \le C \|\boldsymbol{w}_{21}\|_{L^2(\Omega)} \left(\|\overline{\boldsymbol{w}_{21}}\|_{L^2(\Omega)}^{1/2} \|\nabla \overline{\boldsymbol{w}_{21}}\|_{L^2(\Omega)}^{1/2} \|A \overline{\boldsymbol{w}_1}\|_{L^2(\Omega)}\right. \\
&\qquad \left. + \|\nabla \overline{\boldsymbol{w}_1}\|_{L^2(\Omega)}^{1/2} \|A \overline{\boldsymbol{w}_1}\|_{L^2(\Omega)}^{1/2} \|\nabla \overline{\boldsymbol{w}_{21}}\|_{L^2(\Omega)}\right) \\
&\quad = C \|\boldsymbol{w}_{21}\|_{L^2(\Omega)} \left(\|\overline{\boldsymbol{w}_{21}}\|_{L^2(\Omega)}^{1/2} \|\nabla \overline{\boldsymbol{w}_{21}}\|_{L^2(\Omega)}^{1/2} + \|\nabla \overline{\boldsymbol{w}_{21}}\|_{L^2(\Omega)}\right) \\
&\quad \le C\left(1 + \lambda_1^{-1/2}\right) \|\boldsymbol{w}_{21}\|_{L^2(\Omega)} \|\nabla \overline{\boldsymbol{w}_{21}}\|_{L^2(\Omega)} \\
&\quad \le \frac{CC_0}{\nu} \|\nabla \overline{\boldsymbol{w}_{21}}\|_{L^2(\Omega)}^2 + \frac{\nu}{2C_0} \|\boldsymbol{w}_{21}\|_{L^2(\Omega)}^2
\end{aligned} \tag{8.227}
$$

with $C_0 = 2\alpha^2 + \lambda_1^{-1}$. For the last term, one obtains with the definition of $\overline{\boldsymbol{w}_{21}}$, integration by parts, and (8.221)

$$
\begin{aligned}
\frac{\nu}{2C_0} \|\boldsymbol{w}_{21}\|_{L^2(\Omega)}^2 &= \frac{\nu}{2C_0} \left\|\alpha^2 A \overline{\boldsymbol{w}_{21}} + \overline{\boldsymbol{w}_{21}}\right\|_{L^2(\Omega)}^2 \\
&= \frac{\nu}{2C_0}\left(\|\overline{\boldsymbol{w}_{21}}\|_{L^2(\Omega)}^2 + \alpha^4 \|A \overline{\boldsymbol{w}_{21}}\|_{L^2(\Omega)}^2 + 2\alpha^2 \left(A \overline{\boldsymbol{w}_{21}}, \overline{\boldsymbol{w}_{21}}\right)\right) \\
&= \frac{\nu}{2C_0}\left(\frac{1}{\lambda_1} \|\nabla \overline{\boldsymbol{w}_{21}}\|_{L^2(\Omega)}^2 + \alpha^4 \|A \overline{\boldsymbol{w}_{21}}\|_{L^2(\Omega)}^2 + 2\alpha^2 \|\nabla \overline{\boldsymbol{w}_{21}}\|_{L^2(\Omega)}^2\right)
\end{aligned}
$$

$$\leq \frac{\nu}{2C_0}\left(2\alpha^2+\frac{1}{\lambda_1}\right)\left(\|\nabla\,\overline{\boldsymbol{w}_{21}}\,\|^2_{L^2(\Omega)}+\alpha^2\,\|A\,\overline{\boldsymbol{w}_{21}}\,\|^2_{L^2(\Omega)}\right)$$
$$=\frac{\nu}{2}\left(\|\nabla\,\overline{\boldsymbol{w}_{21}}\,\|^2_{L^2(\Omega)}+\alpha^2\,\|A\,\overline{\boldsymbol{w}_{21}}\,\|^2_{L^2(\Omega)}\right).$$

This estimate is inserted in (8.227). Then, this term is absorbed by the left-hand side. Neglecting the arising non-negative term on the left-hand side and neglecting also the dependency of the constant on C_0 and ν, which is not of importance here, one gets

$$\begin{aligned}\frac{d}{dt}\left(\|\,\overline{\boldsymbol{w}_{21}}\,\|^2_{L^2(\Omega)}+\alpha^2\left\|\,\overline{\nabla\boldsymbol{w}_{21}}\,\right\|^2_{L^2(\Omega)}\right) &\leq C\,\|\nabla\,\overline{\boldsymbol{w}_{21}}\,\|^2_{L^2(\Omega)}\\ &= C\alpha^{-2}\alpha^2\,\|\nabla\,\overline{\boldsymbol{w}_{21}}\,\|^2_{L^2(\Omega)}\\ &\leq C\alpha^{-2}\left(\|\,\overline{\boldsymbol{w}_{21}}\,\|^2_{L^2(\Omega)}+\alpha^2\,\|\nabla\,\overline{\boldsymbol{w}_{21}}\,\|^2_{L^2(\Omega)}\right).\end{aligned}$$

Applying Gronwall's lemma, Lemma A.54, yields for almost all $t\in[0,T]$

$$\begin{aligned}&\left(\|\,\overline{\boldsymbol{w}_{21}}\,\|^2_{L^2(\Omega)}+\alpha^2\,\|\nabla\,\overline{\boldsymbol{w}_{21}}\,\|^2_{L^2(\Omega)}\right)(t)\\ &\leq C\left(\|\,\overline{\boldsymbol{w}_{21}}\,\|^2_{L^2(\Omega)}+\alpha^2\,\|\nabla\,\overline{\boldsymbol{w}_{21}}\,\|^2_{L^2(\Omega)}\right)(0)=0,\end{aligned}$$

since $\|\,\overline{\boldsymbol{w}_{21}}\,(0)\|_{L^2(\Omega)}=\|\nabla\,\overline{\boldsymbol{w}_{21}}\,(0)\|_{L^2(\Omega)}=0$ because $\overline{\boldsymbol{w}_1}$ and $\overline{\boldsymbol{w}_2}$ have the same initial data. Therefore, $\overline{\boldsymbol{w}_1}=\overline{\boldsymbol{w}_2}$ in the sense of $H_{\mathrm{div}}(\Omega)$ and V_{div} for almost all $t\in[0,T]$, and with that also $\boldsymbol{w}_1=\boldsymbol{w}_2$. ■

Remark 8.202 (The Hausdorff Dimension of the Global Attractor) In Foias et al. (2002), an estimate for the Hausdorff dimension of the global attractor $\mathcal{A}_{\mathrm{NS}-\alpha}$ of the Navier–Stokes-α model in the case of periodic boundary conditions was derived. Let λ_α be the length scale for which there is a balance of the mean rates of nonlinear transport of energy and viscous dissipation of energy in the Navier–Stokes-α model. This scale depends on α, it is not smaller than the Kolmogorov scale λ, usually it is (much) larger. Then, the estimate proved in Foias et al. (2002) has the form

$$d_{\mathrm{H}}\left(\mathcal{A}_{\mathrm{NS}-\alpha}\right)=\mathcal{O}\left(\left(\frac{L}{\lambda_\alpha}\right)^3\right).$$

Since usually $\lambda_\alpha\gg\lambda$, this dimension is asymptotically smaller than the dimension (8.10) for the Navier–Stokes equations. Besides that one can prove the existence and uniqueness of a weak solution. All these results indicate that the Navier–Stokes-α model is less complex than the Navier–Stokes equations. □

Remark 8.203 (Convergence to a Weak Solution of the Navier–Stokes Equations as $\alpha\to 0$) Similarly as for the Leray-α model, one can show that a subsequence of

$\{\boldsymbol{w}_\alpha\}_{\alpha>0}$ converges to a weak solution of the Navier–Stokes equations as $\alpha \to 0$. The proof of this statement, in the case of periodic boundary conditions, can be found in Foias et al. (2002). In the proof, one shows the uniform (with respect to α) boundedness of some norms from which follows the weak convergence of a subsequence in the corresponding spaces. Further bounds imply, together with Aubin's compactness theorem, e.g., see Foias et al. (2001, Theorem A.11), the strong convergence in appropriate spaces. □

Remark 8.204 (The Finite Element Problem) For sufficiently smooth functions, it is known that from $\nabla \cdot \overline{\boldsymbol{w}}$ in Ω it follows that $\nabla \cdot \boldsymbol{w} = 0$ in Ω, see Remark 8.192. However, finite element functions are not sufficiently smooth, e.g., the strong form of the Laplacian is not well defined, they are only discretely divergence-free, and one has to apply a discrete filter. For these reasons, finite element formulations have been studied in the literature where the discrete divergence constraint was posed for the discrete velocity and the discretely filtered discrete velocity separately.

In Connors (2010), a finite element error analysis for the Navier–Stokes-α model of the following form is presented

$$
\begin{aligned}
\partial_t \boldsymbol{w} - \nu \Delta \boldsymbol{w} + (\nabla \times \boldsymbol{w}) \times \overline{\boldsymbol{w}} + \nabla r = \boldsymbol{f} \quad & \text{in } (0,T] \times \Omega, \\
\nabla \cdot \boldsymbol{w} = 0 \quad & \text{in } (0,T] \times \Omega, \\
\boldsymbol{w} = \boldsymbol{0} \quad & \text{in } (0,T] \times \Gamma, \\
-\alpha^2 \Delta \overline{\boldsymbol{w}} + \overline{\boldsymbol{w}} + \nabla \tilde{r} = \boldsymbol{w} \quad & \text{in } (0,T] \times \Omega, \\
\nabla \cdot \overline{\boldsymbol{w}} = 0 \quad & \text{in } (0,T] \times \Omega, \\
\overline{\boldsymbol{w}} = \boldsymbol{0} \quad & \text{in } (0,T] \times \Gamma, \\
\boldsymbol{w}(0,\cdot) = \boldsymbol{u}_0 \quad & \text{in } \Omega, \\
\int_\Omega r \, d\boldsymbol{x} = \int_\Omega \tilde{r} \, d\boldsymbol{x} = 0 \quad & \text{in } (0,T].
\end{aligned}
\tag{8.228}
$$

Here, Ω is a bounded domain in $\mathbb{R}^d$, $d \in \{2,3\}$, with polyhedral Lipschitz boundary Γ and the form (8.216) of the momentum equation is used. Since there is a separate divergence constraint for the filtered velocity, one needs also a Lagrangian multiplier in the corresponding equation.

The weak form of (8.228) is derived by multiplying the equations with appropriate test functions, integrating the equations in Ω, and applying integration by parts. The correct function spaces with respect to the spatial variable are $V = H_0^1(\Omega)$ and $Q = L_0^2(\Omega)$. In Connors (2010), the continuous-in-time case is considered for conforming finite element spaces: Find $\left(\mathbf{0}symbolw^h, r^h\right) : (0,T] \to V^h \times Q^h$, $\left(\overline{\boldsymbol{w}^h}^h, \tilde{r}^h\right) : (0,T] \to V^h \times Q^h$ with $V^h \subset V$, $Q^h \subset Q$ and

$$
\begin{aligned}
\left(\partial_t \boldsymbol{w}^h, \boldsymbol{v}^h\right) + \left(\nu \nabla \boldsymbol{w}^h, \nabla \boldsymbol{v}^h\right) + n_{\text{rot}}\left(\boldsymbol{w}^h, \overline{\boldsymbol{w}^h}^h, \boldsymbol{v}^h\right) & \\
- \left(\nabla \cdot \boldsymbol{v}^h, r^h\right) = \left(\boldsymbol{f}, \boldsymbol{v}^h\right) \quad & \forall\, \boldsymbol{v}^h \in V^h, \\
\left(\nabla \cdot \boldsymbol{w}^h, q^h\right) = 0 \quad & \forall\, q^h \in Q^h
\end{aligned}
\tag{8.229}
$$

$$\alpha^2 \left(\nabla \overline{\boldsymbol{w}^h}^h, \nabla \tilde{\boldsymbol{v}}^h\right) + \left(\overline{\boldsymbol{w}^h}^h, \tilde{\boldsymbol{v}}^h\right) - \left(\nabla \cdot \tilde{\boldsymbol{v}}^h, \tilde{r}^h\right) = \left(\boldsymbol{w}^h, \tilde{\boldsymbol{v}}^h\right) \quad \forall\ \tilde{\boldsymbol{v}}^h \in V^h,$$

$$\left(\nabla \cdot \overline{\boldsymbol{w}^h}^h, \tilde{q}^h\right) = 0 \quad \forall\ \tilde{q}^h \in Q^h,$$

and $\boldsymbol{w}(0,\cdot) = \boldsymbol{u}_0$. The rotational form of the convective term is defined in (6.21). In addition, a grad-div stabilization term was included in the first equation of (8.229) but not in the equation for the discrete filter in Connors (2010).

The functions $\boldsymbol{w}^h$ and $\overline{\boldsymbol{w}^h}^h$ are contained in V_{div}^h. Considering only velocity test functions from this space, problem (8.229) can be restricted to V_{div}^h: Find $\boldsymbol{w}^h\ :\ (0,T] \to V_{\mathrm{div}}^h$, $\overline{\boldsymbol{w}^h}^h\ :\ (0,T] \to V_{\mathrm{div}}^h$ such that

$$\left(\partial_t \boldsymbol{w}^h, \boldsymbol{v}^h\right) + \left(\nu \nabla \boldsymbol{w}^h, \nabla \boldsymbol{v}^h\right) + n_{\mathrm{rot}}\left(\boldsymbol{w}^h, \overline{\boldsymbol{w}^h}^h, \boldsymbol{v}^h\right) = \left(\boldsymbol{f}, \boldsymbol{v}^h\right), \tag{8.230}$$

$$\alpha^2 \left(\nabla \overline{\boldsymbol{w}^h}^h, \nabla \tilde{\boldsymbol{v}}^h\right) + \left(\overline{\boldsymbol{w}^h}^h, \tilde{\boldsymbol{v}}^h\right) = \left(\boldsymbol{w}^h, \tilde{\boldsymbol{v}}^h\right),$$

for all $\boldsymbol{v}^h, \tilde{\boldsymbol{v}}^h \in V_{\mathrm{div}}^h$. □

Remark 8.205 (Properties of the Discrete Differential Filter) In comparison with the discrete differential filter for the Leray-α model, see (8.196), the discretely filtered finite element velocity for the Navier–Stokes-α model has to satisfy the discrete divergence constraint. Thus, the corresponding equation is defined in V_{div}^h and not in V^h as for the Leray-α model. Inspecting the proofs of Lemmas 8.178–8.180 shows that one can obtain the same results for the discrete differential filter of the Navier–Stokes-α model, where always V^h has to be replaced by V_{div}^h in the formulas and also in the definition of the discrete Laplacian. □

Lemma 8.206 (Existence, Uniqueness, and Stability of the Finite Element Solution) *Let $\boldsymbol{w}_0^h \in V_{\mathrm{div}}^h$ and $\boldsymbol{f} \in L^2(0,t;V')$, then the finite element problem (8.230) possesses a unique solution. The following stability estimate holds for the discretely filtered finite element velocity field*

$$\begin{aligned}
&\left\| \overline{\boldsymbol{w}^h}^h(t) \right\|_{L^2(\Omega)}^2 + \alpha^2 \left\| \nabla \overline{\boldsymbol{w}^h}^h(t) \right\|_{L^2(\Omega)}^2 \\
&\quad + \nu \left(\left\| \nabla \overline{\boldsymbol{w}^h}^h \right\|_{L^2(0,t;L^2(\Omega))}^2 + 2\alpha^2 \left\| \Delta^h \left(\nabla \overline{\boldsymbol{w}^h}^h \right) \right\|_{L^2(0,t;L^2(\Omega))}^2 \right) \\
&\le \left\| \overline{\boldsymbol{w}^h}^h(0) \right\|_{L^2(\Omega)}^2 + \alpha^2 \left\| \nabla \overline{\boldsymbol{w}^h}^h(0) \right\|_{L^2(\Omega)}^2 + \frac{1}{\nu} \left\| \boldsymbol{f} \right\|_{L^2(0,t;V')}^2 .
\end{aligned} \tag{8.231}$$

If $\alpha \le Ch$, then there holds for the unfiltered finite element velocity field

$$\begin{aligned}
&\left\| \boldsymbol{w}^h(t) \right\|_{L^2(\Omega)}^2 + \nu \left\| \nabla \boldsymbol{w}^h \right\|_{L^2(0,t;L^2(\Omega))}^2 \\
&\le C \left(\left\| \boldsymbol{w}^h(0) \right\|_{L^2(\Omega)}^2 + \alpha^2 \left\| \nabla \boldsymbol{w}^h(0) \right\|_{L^2(\Omega)}^2 + \frac{1}{\nu} \left\| \boldsymbol{f} \right\|_{L^2(0,t;V')}^2 \right),
\end{aligned} \tag{8.232}$$

where the constant depends on the constant of the inverse estimate (C.37) *but not on the mesh width h.*

Proof As usual, stability estimates are derived by using appropriate test functions in the equation. Considering (8.230), then the test function has to be chosen such that the nonlinear convective term vanishes, i.e., one has to choose $\boldsymbol{v}^h = \overline{\boldsymbol{w}^h}^h$ which gives the desired result, see (6.26), and leads to the equation

$$\left(\partial_t \boldsymbol{w}^h, \overline{\boldsymbol{w}^h}^h\right) + \left(\nu \nabla \boldsymbol{w}^h, \nabla \overline{\boldsymbol{w}^h}^h\right) = \left(\boldsymbol{f}, \overline{\boldsymbol{w}^h}^h\right). \tag{8.233}$$

With property (8.197) and the commutation of filtering and temporal derivative, compare Lemma 8.197, one obtains

$$\left(\partial_t \boldsymbol{w}^h, \overline{\boldsymbol{w}^h}^h\right) = \left(\overline{\partial_t \boldsymbol{w}^h}^h, \boldsymbol{w}^h\right) = \left(\partial_t \overline{\boldsymbol{w}^h}^h, \boldsymbol{w}^h\right),$$

from what

$$\left(\partial_t \boldsymbol{w}^h, \overline{\boldsymbol{w}^h}^h\right) = \frac{1}{2}\left(\partial_t \boldsymbol{w}^h, \overline{\boldsymbol{w}^h}^h\right) + \frac{1}{2}\left(\boldsymbol{w}^h, \partial_t \overline{\boldsymbol{w}^h}^h\right)$$

follows. Then, the product rule shows that

$$\frac{1}{2}\frac{d}{dt}\left(\boldsymbol{w}^h, \overline{\boldsymbol{w}^h}^h\right) = \frac{1}{2}\left(\partial_t \boldsymbol{w}^h, \overline{\boldsymbol{w}^h}^h\right) + \frac{1}{2}\left(\boldsymbol{w}^h, \partial_t \overline{\boldsymbol{w}^h}^h\right) = \left(\partial_t \boldsymbol{w}^h, \overline{\boldsymbol{w}^h}^h\right).$$

Inserting this expression in (8.233) gives

$$\frac{1}{2}\frac{d}{dt}\left(\boldsymbol{w}^h, \overline{\boldsymbol{w}^h}^h\right) + \left(\nu \nabla \boldsymbol{w}^h, \nabla \overline{\boldsymbol{w}^h}^h\right) = \left(\boldsymbol{f}, \overline{\boldsymbol{w}^h}^h\right). \tag{8.234}$$

Using now the definition of the discrete filter (8.230) and (8.201) on the left-hand side, and the estimate for the duality pairing and Young's inequality (A.5) on the right-hand side yields

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\left(\left\|\overline{\boldsymbol{w}^h}^h\right\|_{L^2(\Omega)}^2 + \alpha^2\left\|\nabla \overline{\boldsymbol{w}^h}^h\right\|_{L^2(\Omega)}^2\right) \\
&+\nu\left(\left\|\nabla \overline{\boldsymbol{w}^h}^h\right\|_{L^2(\Omega)}^2 + \alpha^2\left\|\Delta^h\left(\nabla \overline{\boldsymbol{w}^h}^h\right)\right\|_{L^2(\Omega)}^2\right) \le \|\boldsymbol{f}\|_{V'}\left\|\nabla \overline{\boldsymbol{w}^h}^h\right\|_{L^2(\Omega)} \\
&\qquad\qquad \le \frac{1}{2\nu}\|\boldsymbol{f}\|_{V'}^2 + \frac{\nu}{2}\left\|\nabla \overline{\boldsymbol{w}^h}^h\right\|_{L^2(\Omega)}^2.
\end{aligned}$$

The last term can be absorbed from the left-hand side. Multiplying the resulting estimate by two and integrating in $(0, T)$ gives (8.231).

To prove (8.232), norms of $\overline{\boldsymbol{w}^h}^h$ will be bounded from below with norms of $\boldsymbol{w}^h$. Inserting $\tilde{\boldsymbol{v}}^h = \boldsymbol{w}^h \in V_{\mathrm{div}}^h$ in the definition of the discrete filter (8.230), applying the Cauchy–Schwarz inequality (A.10) and the inverse inequality (C.37) gives

$$\begin{aligned}
\left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)}^2 &= \left(\overline{\boldsymbol{w}^h}^h, \boldsymbol{w}^h\right) + \alpha^2 \left(\nabla \overline{\boldsymbol{w}^h}^h, \nabla \boldsymbol{w}^h\right) \\
&\le \left\|\overline{\boldsymbol{w}^h}^h\right\|_{L^2(\Omega)} \left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)} + \alpha^2 \left\|\nabla \overline{\boldsymbol{w}^h}^h\right\|_{L^2(\Omega)} \left\|\nabla \boldsymbol{w}^h\right\|_{L^2(\Omega)} \\
&\le \left\|\overline{\boldsymbol{w}^h}^h\right\|_{L^2(\Omega)} \left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)} + \frac{\alpha^2 C_{\mathrm{inv}}^2}{h^2} \left\|\overline{\boldsymbol{w}^h}^h\right\|_{L^2(\Omega)} \left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)}.
\end{aligned}$$

Using the condition on α gives

$$\left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)}^2 \le C \left\|\overline{\boldsymbol{w}^h}^h\right\|_{L^2(\Omega)}. \tag{8.235}$$

Choosing in (8.230) $\tilde{\boldsymbol{v}}^h$ to be the discrete Laplacian $\Delta^h \boldsymbol{w}^h \in V_{\mathrm{div}}^h$, see Remark 8.205, and applying the definition (8.200) of the discrete Laplacian yields with the same tools the estimate

$$\begin{aligned}
\left\|\nabla \boldsymbol{w}^h\right\|_{L^2(\Omega)}^2 &= \left(\nabla \overline{\boldsymbol{w}^h}^h, \nabla \boldsymbol{w}^h\right) + \alpha^2 \left(\Delta^h \overline{\boldsymbol{w}^h}^h, \Delta^h \boldsymbol{w}^h\right) \\
&\le \left\|\nabla \overline{\boldsymbol{w}^h}^h\right\|_{L^2(\Omega)} \left\|\nabla \boldsymbol{w}^h\right\|_{L^2(\Omega)} + \alpha^2 \left\|\Delta^h \overline{\boldsymbol{w}^h}^h\right\|_{L^2(\Omega)} \left\|\Delta^h \boldsymbol{w}^h\right\|_{L^2(\Omega)}.
\end{aligned} \tag{8.236}$$

Using in the definition of the discrete Laplacian as test function the discrete Laplacian itself and applying the inverse inequality gives

$$\left\|\Delta^h \boldsymbol{w}^h\right\|_{L^2(\Omega)}^2 \le \left\|\nabla \boldsymbol{w}^h\right\|_{L^2(\Omega)} \left\|\nabla \Delta^h \boldsymbol{w}^h\right\|_{L^2(\Omega)} \le \frac{C_{\mathrm{inv}}}{h} \left\|\nabla \boldsymbol{w}^h\right\|_{L^2(\Omega)} \left\|\Delta^h \boldsymbol{w}^h\right\|_{L^2(\Omega)},$$

which gives an inverse estimate for the discrete Laplacian. Inserting this inverse estimate in (8.236) and using the assumption on α yields

$$\left\|\nabla \boldsymbol{w}^h\right\|_{L^2(\Omega)}^2 \le C \left\|\nabla \overline{\boldsymbol{w}^h}^h\right\|_{L^2(\Omega)}. \tag{8.237}$$

Now, (8.232) is obtained by neglecting the terms with α on the left-hand side of (8.231), estimating the other terms on the left-hand side with (8.235) and (8.237), and bounding the terms from the initial condition with (8.198).

Estimate (8.232) for $\boldsymbol{w}^h$ is of the same form as estimate (7.11). In addition, by using (8.235), (8.231), and (8.198), an estimate of form (7.10) can be derived for $\boldsymbol{w}^h$. Then, the existence and uniqueness of a velocity solution $\boldsymbol{w}^h \in L^2\left(0, T; V_{\mathrm{div}}^h\right) \cap L^\infty\left(0, T; L^2(\Omega)\right)$ can be proved along the lines of the proof of Lemma 7.12. The existence and uniqueness of $\overline{\boldsymbol{w}^h}^h$ follows from the unique solvability of the

Helmholtz equation in (8.230), which can be proved with the Theorem of Lax–Milgram, see Theorem B.4. ■

Remark 8.207 (Further Results From Finite Element Analysis)

- An error estimate for the continuous-in-time case can be found in Connors (2010). The proof of this estimate follows the standard lines, e.g., as the proof of Theorem 7.35 for the Navier–Stokes equations. It uses some properties of the discrete differential filter which were also used in the proof of Lemma 8.206. The nonlinear convective term is estimated with inequalities that can be found in Sect. 6.1.2. Under some regularity assumptions, one obtains an estimate for

$$\left\| \left(\boldsymbol{u} - \boldsymbol{w}^h\right)(t) \right\|_{L^2(\Omega)}^2 + \nu \left\| \nabla \left(\boldsymbol{u} - \boldsymbol{w}^h\right) \right\|_{L^2(0,t;L^2(\Omega))}^2$$

  with a factor of size $\exp(C\nu^{-3}t)$, similarly to the Navier–Stokes equations, see (7.44), and the Leray-α model, see (8.209), (8.211). It is assumed that $\alpha = \mathcal{O}(h)$, which leads to second order convergence, as for the Leray-α model, see Remark 8.185.
- A full discretization with the Crank–Nicolson scheme as time integrator was considered in Layton et al. (2010). In this paper, the stability of the finite element solution is proved, but an error analysis is not presented.
- An error analysis of a fully discrete scheme was presented in Miles and Rebholz (2010). The studied scheme uses the Crank–Nicolson method and the rotational form (8.217). A key feature consists in enforcing that the curl of the solution $\boldsymbol{w}^h$ should be discretely divergence-free. With this requirement, improved conservation properties of the scheme can be proved. In addition, the analysis presented in Miles and Rebholz (2010) includes a stability estimate, the proof of the existence of a solution, and a finite element error estimate. Furthermore, a Navier–Stokes-α deconvolution model is presented, where $\overline{\boldsymbol{w}} \times (\nabla \times \boldsymbol{w})$ is replaced with $G_{\text{deconv},N}(\overline{\boldsymbol{w}}) \times (\nabla \times \boldsymbol{w})$ and $G_{\text{deconv},N}$ being the van Cittert approximate deconvolution operator, compare Remark 8.156.

□

Example 8.208 (Turbulent Channel Flow at $Re_\tau = 180$) The setup of this example was exactly as in Examples 8.128 and 8.188. The Navier–Stokes-α model was applied as well in the convective form (8.215) and also in the rotational form (8.216) with $\alpha = \alpha_0 h_{K,\text{short}}$.

It turned out that the simulations blew up if α_0 was chosen too small for both formulations. That solving the nonlinear problem arising in the Navier–Stokes-α model was not possible is also reported for simulations of a two-dimensional flow around a cylinder, Example D.9, in Layton et al. (2010).

Results obtained for the smallest values of α_0 for which there was no blow-up are presented in Fig. 8.10. It can be observed that not even the mean velocity field was reproduced reasonably accurately for these values.

In Abdi (2015), the Navier–Stokes-α model (8.215) was tested without the zeroth order term, such that the only difference to the Leray-α model is that $\overline{\boldsymbol{w}}$ is discretely

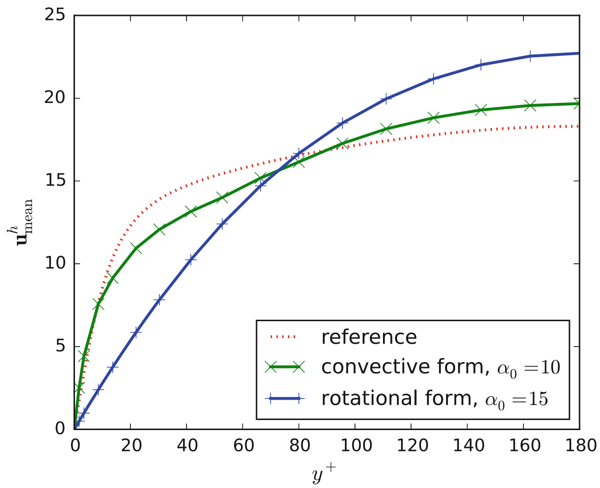

Fig. 8.10 Example 8.208. Mean velocity field obtained for the Navier–Stokes-α model with different values of $\alpha = \alpha_0 h_{K,\text{short}}$

divergence-free. Without zeroth order term, stable simulations were observed and reasonably accurate results for the turbulent channel flow at $\text{Re}_\tau = 180$ were obtained. Thus, the instability arises from the zeroth order term.

A reduced Navier–Stokes-α model is proposed and analyzed in Cuff et al. (2015), leading to stable simulations with good results for the mean velocity profile of the turbulent channel flow problem at $\text{Re}_\tau = 180$. In this model, the equation defining $\overline{\boldsymbol{w}}$ in (8.215) is inserted in the momentum balance (8.216) with the rotational form of the convective term. The arising fourth order terms are approximated by second order terms with the help of the van Cittert approximate deconvolution, see Remark 8.156. In the first step, the term $-\alpha^2 \Delta\,(\partial_t \overline{\boldsymbol{w}}\,)$ arises that provides additional stability since it is a viscous term. A similar approach was used for deriving a reduced ADM, compare Remark 8.164. □

8.8 Variational Multiscale Methods

Remark 8.209 (Contents) This section presents so-called Variational Multiscale (VMS) methods. Common features of these methods comprise that they are based on a variational formulation of the Navier–Stokes equations and that the scales are defined by projection into function spaces. Apart of these features, realizations of VMS methods are quite different. However, the proposed methods can be divided into two classes: two-scale and three-scale VMS methods, see Sect. 8.8.1.

A recent review on VMS methods, which is in some parts similar to the presentation in this section, can be found in Ahmed et al. (2016). In this review, two more VMS methods than presented here are included, the LPS method, see Sect. 5.4 starting with Remarks 5.52, and the higher order term-by-term stabilization method, see Remark 5.57. A longer presentation of LPS methods can be found in (Roos et al.

2008, Chap. IV.4). The link of the LPS method to VMS methods was established in Braack and Burman (2006). Additionally, considerably more information on the numerical experience with some methods is provided in Ahmed et al. (2016). □

8.8.1 *Basic Concepts*

Remark 8.210 (Differences to LES) Similarly to classical LES methods, Variational Multiscale (VMS) methods seek to simulate only large flow structures. Therefore, these methods are also called VMS-LES. However, there are some fundamental differences between VMS methods and classical LES methods.

The difficulties of the classical LES methods in their mathematical study originate in the definition of the large scales by spatial averaging. As an alternative, VMS methods consider large scales that are defined by projection into appropriate spaces. To this end, a variational formulation of the Navier–Stokes equations is considered, again in contrast to LES methods, whose derivation is based on the strong form of the Navier–Stokes equations, see Sects. 8.2.2 and 8.2.3. The consideration of a variational form of the equation and the use of projections for defining the different scales allow to incorporate bounded domains and boundary conditions into the mathematical analysis in a natural way.

There are VMS methods that decompose the flow field into two scales, resolved and unresolved ones, like LES methods, see Remark 8.23. However, the VMS methodology allows also the decomposition into more than two scales. An alternative approach are so-called three-scale VMS methods. For these methods, there is another difference to LES methods: the turbulence model does not act directly on all resolved scales but only on the smallest resolved scales, see Remark 8.214.

First ideas of projection-based methods, also for different problems than the Navier–Stokes equations, can be found in Hughes (1995), Guermond (1999a), and Hughes et al. (2000). □

8.8.1.1 Two-Scale VMS Methods

Remark 8.211 (Basic Approach for a Two-Scale VMS Method) A two-scale VMS method uses just a decomposition of the scales in resolved scales $(\overline{\boldsymbol{u}},\overline{p})$ and unresolved scales $(\boldsymbol{u}',p')$ with

$$\boldsymbol{u} = \overline{\boldsymbol{u}} + \boldsymbol{u}', \quad p = \overline{p} + p'.$$

Inserting this decomposition in (8.243) and using the same decomposition for the test functions yields

$$A\left(\boldsymbol{u};(\overline{\boldsymbol{u}},\overline{p}),(\overline{\boldsymbol{v}},\overline{q})\right) + A\left(\boldsymbol{u};\left(\boldsymbol{u}',p'\right),(\overline{\boldsymbol{v}},\overline{q})\right) = F\left(\overline{\boldsymbol{v}}\right), \tag{8.238}$$

$$A\left(\boldsymbol{u};(\overline{\boldsymbol{u}},\overline{p}),\left(\boldsymbol{v}',q'\right)\right) + A\left(\boldsymbol{u};\left(\boldsymbol{u}',p'\right),\left(\boldsymbol{v}',q'\right)\right) = F\left(\boldsymbol{v}'\right). \tag{8.239}$$

To simplify notations, set

$$\boldsymbol{U} = \begin{pmatrix} \boldsymbol{u} \\ p \end{pmatrix}, \quad \boldsymbol{V} = \begin{pmatrix} \boldsymbol{v} \\ q \end{pmatrix} \quad \text{and so on.}$$

Decomposing the form $A\left(\cdot;\cdot,\cdot\right)$ into its linear part and the trilinear convective term

$$A\left(\boldsymbol{u};\boldsymbol{U},\boldsymbol{V}\right) = A_{\text{lin}}\left(\boldsymbol{U},\boldsymbol{V}\right) + n(\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}),$$

Eq. (8.239) can be written in the form

$$A_{\boldsymbol{U}}\left(\boldsymbol{U}',\boldsymbol{V}'\right) + n\left(\boldsymbol{u}',\boldsymbol{u}',\boldsymbol{v}'\right) = -\left\langle \mathbf{Res}\left(\overline{\boldsymbol{U}}\right), \boldsymbol{V}'\right\rangle_{(V\times Q)',(V\times Q)} \tag{8.240}$$

with

$$\begin{aligned} A_{\boldsymbol{U}}\left(\boldsymbol{U}',\boldsymbol{V}'\right) &= A_{\text{lin}}\left(\boldsymbol{U}',\boldsymbol{V}'\right) + n\left(\boldsymbol{u}',\overline{\boldsymbol{u}},\boldsymbol{v}'\right) + n\left(\overline{\boldsymbol{u}},\boldsymbol{u}',\boldsymbol{v}'\right), \\ \left\langle \mathbf{Res}\left(\overline{\boldsymbol{U}}\right), \boldsymbol{V}'\right\rangle_{(V\times Q)',(V\times Q)} &= A_{\text{lin}}\left(\overline{\boldsymbol{U}},\boldsymbol{V}'\right) + n\left(\overline{\boldsymbol{u}},\overline{\boldsymbol{u}},\boldsymbol{v}'\right) - \langle \boldsymbol{f},\boldsymbol{v}'\rangle_{V',V}. \end{aligned}$$

A possible interpretation of (8.240) is that the unresolved scales are a function of the residual of the resolved scales. Then, they can be represented in the form

$$\boldsymbol{U}' = \boldsymbol{F}_{\boldsymbol{U}}\left(-\mathbf{Res}\left(\overline{\boldsymbol{U}}\right)\right). \tag{8.241}$$

Inserting this representation in (8.238) gives an equation for the resolved scales, see Sect. 8.8.2 for a concrete method.

There is a second way to interprete (8.240). Taking into account that the left-hand side of (8.240) depends on the temporal derivative of the small scale velocity, it is possible to model the unresolved scales as a function of the residual of the large scales and the small scale velocity $\boldsymbol{u}'_{\text{old}}$ at former times

$$\boldsymbol{U}' = \boldsymbol{F}_{\boldsymbol{U}}\left(-\mathbf{Res}\left(\overline{\boldsymbol{U}}\right), \boldsymbol{u}'_{\text{old}}\right). \tag{8.242}$$

This interpretation is used in the derivation of the method presented in Sect. 8.8.3.

In practice, the operator $\boldsymbol{F}_{\boldsymbol{U}}$ is not known. Even if it would be available, using this operator for modeling (representing) the unresolved scales would not lead to a less complex problem than the Navier–Stokes equations, i.e., there would be no turbulence modeling. A two-scale VMS method aims to approximate $\boldsymbol{F}_{\boldsymbol{U}}$, see Sects. 8.8.2 and 8.8.3 for possible approaches. □

Remark 8.212 (On $A_{\boldsymbol{U}}\left(\boldsymbol{U}',\boldsymbol{V}'\right)$) The operator $A_{\boldsymbol{U}}\left(\boldsymbol{U}',\boldsymbol{V}'\right)$ is the Gâteaux derivative of $A(\cdot;\cdot,\cdot)$ at $\boldsymbol{U}$ in the direction of $\boldsymbol{U}'$ since one obtains with the linearity of $A_{\text{lin}}\left(\cdot,\cdot\right)$

and the trilinearity of $n(\cdot,\cdot,\cdot)$

$$\lim_{\varepsilon\to 0} \frac{A\left(\overline{\boldsymbol{u}}+\varepsilon\boldsymbol{u}';\overline{\boldsymbol{U}}+\varepsilon\boldsymbol{U}',\boldsymbol{V}'\right)-A\left(\overline{\boldsymbol{u}};\overline{\boldsymbol{U}},\boldsymbol{V}'\right)}{\varepsilon}$$
$$= \lim_{\varepsilon\to 0}\Big(A_{\text{lin}}\left(\overline{\boldsymbol{U}}+\varepsilon\boldsymbol{U}',\boldsymbol{V}'\right)+n\left(\overline{\boldsymbol{u}}+\varepsilon\boldsymbol{u}',\overline{\boldsymbol{u}}+\varepsilon\boldsymbol{u}',\boldsymbol{v}'\right)-A_{\text{lin}}\left(\overline{\boldsymbol{U}},\boldsymbol{V}'\right)$$
$$-n\left(\overline{\boldsymbol{u}},\overline{\boldsymbol{u}},\boldsymbol{v}'\right)\Big)/\varepsilon$$
$$= A_{\text{lin}}\left(\boldsymbol{U}',\boldsymbol{V}'\right)+n\left(\boldsymbol{u}',\overline{\boldsymbol{u}},\boldsymbol{v}'\right)+n\left(\overline{\boldsymbol{u}},\boldsymbol{u}',\boldsymbol{v}'\right).$$

□

Remark 8.213 (Turbulence Modeling Without Physical Turbulence Model) The two-scale VMS approach usually does not include a turbulence model that is based on physical considerations, like the Smagorinsky model. The turbulence modeling is based on purely mathematical considerations by defining the operator $\boldsymbol{F_U}$ in (8.241) or (8.242). Note that the arguments of this operator are in practice quantities that are computed in the simulations.

It is shown in Guasch and Codina (2013), for the quasi-static OSS-VMS method, see Remark 8.234, and with several simplifying assumptions that nevertheless this kind of mathematical based turbulence modeling may have the correct physical behavior in the inertial subrange of turbulent flows. Concretely, it is shown, with heuristic arguments, that the rate of transfer of energy contributed from the model of the unresolved scales does not depend on the mesh width and it is proportional to the mean molecular dissipation rate, see Remark 8.8. □

8.8.1.2 Three-Scale VMS Methods

Remark 8.214 (Basic Approach for a Three-Scale VMS Method) Consider the Navier–Stokes equations (7.1) in a bounded domain, equipped for simplicity with no-slip conditions, and a decomposition of the flow into three scales: the large scales $(\overline{\boldsymbol{u}},\overline{p})$, the small resolved scales $(\hat{\boldsymbol{u}},\hat{p})$, and the unresolved scales $(\boldsymbol{u}',p')$, with

$$\boldsymbol{u}=\overline{\boldsymbol{u}}+\hat{\boldsymbol{u}}+\boldsymbol{u}',\quad p=\overline{p}+\hat{p}+p'.$$

The starting point of a VMS method is the variational formulation of the Navier–Stokes equations, e.g., like (7.40). This formulation can be written in short form

$$A\left(\boldsymbol{u};(\boldsymbol{u},p),(\boldsymbol{v},q)\right)=F\left(\boldsymbol{v}\right). \tag{8.243}$$

Decomposing the test functions also into three scales, the variational form of the Navier–Stokes equations can be written as a coupled system: Find $\boldsymbol{u}=\overline{\boldsymbol{u}}+\hat{\boldsymbol{u}}+\boldsymbol{u}'$: $(0,T]\to V$, $p=\overline{p}+\hat{p}+p'$: $(0,T]\to Q$ satisfying for all $(\boldsymbol{v},q)\in V\times Q$ with

$\boldsymbol{v} = \overline{\boldsymbol{v}} + \hat{\boldsymbol{v}} + \boldsymbol{v}', q = \overline{q} + \hat{q} + q'$

$$A\left(\boldsymbol{u};(\overline{\boldsymbol{u}},\overline{p}),(\overline{\boldsymbol{v}},\overline{q})\right) + A\left(\boldsymbol{u};(\hat{\boldsymbol{u}},\hat{p}),(\overline{\boldsymbol{v}},\overline{q})\right) + A\left(\boldsymbol{u};\left(\boldsymbol{u}',p'\right),(\overline{\boldsymbol{v}},\overline{q})\right) = F\left(\overline{\boldsymbol{v}}\right),$$
$$A\left(\boldsymbol{u};(\overline{\boldsymbol{u}},\overline{p}),(\hat{\boldsymbol{v}},\hat{q})\right) + A\left(\boldsymbol{u};(\hat{\boldsymbol{u}},\hat{p}),(\hat{\boldsymbol{v}},\hat{q})\right) + A\left(\boldsymbol{u};\left(\boldsymbol{u}',p'\right),(\hat{\boldsymbol{v}},\hat{q})\right) = F\left(\hat{\boldsymbol{v}}\right),$$
$$A\left(\boldsymbol{u};(\overline{\boldsymbol{u}},\overline{p}),\left(\boldsymbol{v}',q'\right)\right) + A\left(\boldsymbol{u};(\hat{\boldsymbol{u}},\hat{p}),\left(\boldsymbol{v}',q'\right)\right) + A\left(\boldsymbol{u};\left(\boldsymbol{u}',p'\right),\left(\boldsymbol{v}',q'\right)\right) = F\left(\boldsymbol{v}'\right).$$

Here, the linearity of the variational problem with respect to the test function has been used. Now, the basic ideas and assumptions of a three-scale VMS method are as follows:

- The equation with the test function from the unresolved scales is neglected, i.e., the equations are projected in the space of the resolved scales.
- It is assumed that the direct influence of the unresolved scales on the large scales is negligible, i.e., $A\left(\boldsymbol{u};(\boldsymbol{u}',p'),(\overline{\boldsymbol{v}},\overline{q})\right) = 0$. The unresolved scales might influence the resolved scales, e.g., by a backscatter of energy. It was mentioned in Remark 8.7 that this backscatter occurs mainly to the next larger eddies, which are represented by the small resolved scales. In this sense, this assumption is reasonable.
- The influence of the unresolved scales onto the small resolved scales is modeled:

$$A\left(\boldsymbol{u};\left(\boldsymbol{u}',p'\right),(\hat{\boldsymbol{v}},\hat{q})\right) \approx T\left(\boldsymbol{u};(\overline{\boldsymbol{u}},\overline{p}),(\hat{\boldsymbol{u}},\hat{p}),(\hat{\boldsymbol{v}},\hat{q})\right).$$

  The choice of the model $T\left(\boldsymbol{u};(\overline{\boldsymbol{u}},\overline{p}),(\hat{\boldsymbol{u}},\hat{p}),(\hat{\boldsymbol{v}},\hat{q})\right)$ may be guided by physical ideas in turbulence modeling, e.g., eddy viscosity models of Smagorinsky type (8.67), for $\overline{\boldsymbol{u}}$ or $\hat{\boldsymbol{u}}$, are often used. From the numerical point of view, the turbulence model $T\left(\boldsymbol{u};(\overline{\boldsymbol{u}},\overline{p}),(\hat{\boldsymbol{u}},\hat{p}),(\hat{\boldsymbol{v}},\hat{q})\right)$ introduces additional viscosity that acts as stabilization.

Let $\overline{V}, \overline{Q}$ be spaces representing the large scales and $\widehat{V}, \widehat{Q}$ spaces for the resolved small scales. An abstract three-scale VMS method reads as a coupled system of the form: Find $(\overline{\boldsymbol{u}},\hat{\boldsymbol{u}},\overline{p},\hat{p}) : (0,T] \to \overline{V} \times \widehat{V} \times \overline{Q} \times \widehat{Q}$ such that

$$A\left(\overline{\boldsymbol{u}}+\hat{\boldsymbol{u}};(\overline{\boldsymbol{u}},\overline{p}),(\overline{\boldsymbol{v}},\overline{q})\right) + A\left(\overline{\boldsymbol{u}}+\hat{\boldsymbol{u}};(\hat{\boldsymbol{u}},\hat{p}),(\overline{\boldsymbol{v}},\overline{q})\right) = F\left(\overline{\boldsymbol{v}}\right),$$
$$A\left(\overline{\boldsymbol{u}}+\hat{\boldsymbol{u}};(\overline{\boldsymbol{u}},\overline{p}),(\hat{\boldsymbol{v}},\hat{q})\right) + A\left(\overline{\boldsymbol{u}}+\hat{\boldsymbol{u}};(\hat{\boldsymbol{u}},\hat{p}),(\hat{\boldsymbol{v}},\hat{q})\right) \tag{8.244}$$
$$+T\left(\overline{\boldsymbol{u}}+\hat{\boldsymbol{u}};(\overline{\boldsymbol{u}},\overline{p}),(\hat{\boldsymbol{u}},\hat{p}),(\hat{\boldsymbol{v}},\hat{q})\right) = F\left(\hat{\boldsymbol{v}}\right) \tag{8.245}$$

for all $(\overline{\boldsymbol{v}},\hat{\boldsymbol{v}},\overline{q},\hat{q}) \in \overline{V} \times \widehat{V} \times \overline{Q} \times \widehat{Q}$.

Note that a characteristic feature of a three-scale VMS method is that the model for the influence of the unresolved scales acts directly only on the small resolved scales. Since the small resolved scales and the large scales are coupled in (8.244), (8.245), the model $T\left(\boldsymbol{u};(\overline{\boldsymbol{u}},\overline{p}),(\hat{\boldsymbol{u}},\hat{p}),(\hat{\boldsymbol{v}},\hat{q})\right)$ influences the large scales indirectly. This situation is in contrast to classical LES models, where the model acts directly on all resolved scales.

To specify a concrete three-scale VMS method, one has to define the spaces $\overline{V}, \widehat{V}, \overline{Q}, \widehat{Q}$ and a model $T\left(\overline{\boldsymbol{u}}+\hat{\boldsymbol{u}};(\overline{\boldsymbol{u}}, \overline{p}),(\hat{\boldsymbol{u}}, \hat{p}),(\hat{\boldsymbol{v}}, \hat{q})\right)$. □

Remark 8.215 (Principal Approaches for Choosing Appropriate Spaces in Three-Scale VMS Methods) Concerning finite element methods for discretizing (8.244), (8.245), there are at least two different approaches.

Standard finite element spaces can be used for the large scales $\overline{V} \times \overline{Q}$. The finite element spaces $\widehat{V} \times \widehat{Q}$ need to have a higher resolution since they should represent smaller scales. A proposal consists in using mesh cell bubble functions for this purpose, see Sect. 8.8.4 for details.

The second way for choosing the spaces consists in using a common standard finite element space for all resolved scales and an additional large scale space. Methods of this type will be presented in Sects. 8.8.5 and 8.8.6. □

Remark 8.216 (Common Goal of Three-Scale VMS Methods and the Dynamic Smagorinsky Model) Using for $T\left(\overline{\boldsymbol{u}}+\hat{\boldsymbol{u}};(\overline{\boldsymbol{u}}, \overline{p}),(\hat{\boldsymbol{u}}, \hat{p}),(\hat{\boldsymbol{v}}, \hat{q})\right)$ a model of Smagorinsky-type, as it is often done, then the principle goal of a three-scale VMS method and the dynamic Smagorinsky model presented in Remark 8.126 is similar. Based on the experience that the use of the Smagorinsky model (8.67) with a fixed constant C_S as turbulence model introduces too much viscosity, one tries to reduce the influence of this model in accordance to the local flow field. In the dynamic Smagorinsky model, this reduction is done by using a function $C_S(t, \boldsymbol{x})$ and adjusting this function appropriately. In a three-scale VMS method, the reduction is performed by choosing an appropriate small resolved scale space to which the direct influence of the Smagorinsky model is restricted. □

Remark 8.217 (Notation) To be consistent with the presentation of the other turbulence models, the large scales computed with a VMS method are denoted by $(\boldsymbol{w}^h, r^h)$ and the small resolved scales by $\left(\hat{\boldsymbol{w}}^h, \hat{r}^h\right)$. □

8.8.2 A Two-Scale Residual-Based VMS Method

Remark 8.218 (Contents) This section presents a two-scale VMS method that was proposed in Bazilevs et al. (2007). The derivation of this method is based on a perturbation series with respect to the norm of the residual of the resolved scales. Truncating this series after the first term and applying some modeling of this term leads finally to a turbulence model which can be considered as an extension of the SUPG/PSPG/grad-div stabilization presented in Sect. 5.3.2. □

Remark 8.219 (Scale Separation by Projection) It will be assumed that the decomposition of the spaces for resolved and unresolved scales is of the form

$$V \times Q = \left(\overline{V} \oplus V'\right) \times \left(\overline{Q} \oplus Q'\right),$$

where the scales are defined by a projection. For instance, let $\boldsymbol{v} \in V$, then possibilities are the $L^2(\Omega)$ projection $\overline{\boldsymbol{v}} = P_{L^2}\boldsymbol{v}$ or the elliptic projection $\overline{\boldsymbol{v}} = P_{H^1}\boldsymbol{v}$.

The decomposition of V into a direct sum induces the assumption that the resolved velocity scales and the unresolved velocity scales possess also homogeneous Dirichlet boundary conditions as the functions from V. □

Remark 8.220 (The Perturbation Series) The derivation of the two-scale residual-based VMS method is based on a perturbation series for a potentially small quantity. This quantity is $\varepsilon = \left\|\mathbf{Res}\left(\overline{\boldsymbol{U}}\right)\right\|_{(V\times Q)'}$. For this quantity to be small, the space $\overline{V}\times\overline{Q}$ has to be sufficiently large. In fact, it is assumed that the larger $\overline{V}\times\overline{Q}$, the better $\overline{\boldsymbol{U}}$ approximates $\boldsymbol{U}$ and the smaller is $\mathbf{Res}\left(\overline{\boldsymbol{U}}\right)$. Then, the perturbation series is of the form

$$\boldsymbol{U}' = \varepsilon\boldsymbol{U}_1' + \varepsilon^2\boldsymbol{U}_2' + \ldots = \sum_{i=1}^{\infty}\varepsilon^i\boldsymbol{U}_i'. \tag{8.246}$$

Thus, this approach induces that if $\varepsilon = 0$, i.e., $\mathbf{Res}\left(\overline{\boldsymbol{U}}\right) = \mathbf{0}$, then $\boldsymbol{U}' = \boldsymbol{F}_U\left(-\mathbf{Res}\left(\overline{\boldsymbol{U}}\right)\right) = \mathbf{0}$. □

Remark 8.221 (The Equation for the Unresolved Scales with Perturbation Series) Inserting the perturbation series (8.246) in the terms of Eq. (8.240) for the unresolved scales and using the linearity of the forms in the respective arguments yields

$$A_{\boldsymbol{U}}\left(\sum_{i=1}^{\infty}\varepsilon^i\boldsymbol{U}_i', \boldsymbol{V}'\right) = \sum_{i=1}^{\infty}\varepsilon^i A_{\boldsymbol{U}}\left(\boldsymbol{U}_i', \boldsymbol{V}'\right)$$

and

$$\begin{aligned}
& n\left(\sum_{i=1}^{\infty}\varepsilon^i\boldsymbol{u}_i', \sum_{i=1}^{\infty}\varepsilon^i\boldsymbol{u}_i', \boldsymbol{v}'\right) \\
&= \varepsilon^2 n\left(\boldsymbol{u}_1', \boldsymbol{u}_1', \boldsymbol{v}'\right) + \varepsilon^3 n\left(\boldsymbol{u}_1', \boldsymbol{u}_2', \boldsymbol{v}'\right) + n\left(\boldsymbol{u}_2', \boldsymbol{u}_1', \boldsymbol{v}'\right) + \ldots \\
&= \sum_{i=2}^{\infty}\varepsilon^i\left(\sum_{j=1}^{i-1} n\left(\boldsymbol{u}_i', \boldsymbol{u}_j', \boldsymbol{v}'\right)\right).
\end{aligned}$$

These terms can be inserted in (8.240) leading to

$$\begin{aligned}
& \sum_{i=1}^{\infty}\varepsilon^i A_{\boldsymbol{U}}\left(\boldsymbol{U}_i', \boldsymbol{V}'\right) + \sum_{i=2}^{\infty}\varepsilon^i\left(\sum_{j=1}^{i-1} n\left(\boldsymbol{u}_i', \boldsymbol{u}_j', \boldsymbol{v}'\right)\right) \\
&= -\varepsilon\left\langle \frac{\mathbf{Res}\left(\overline{\boldsymbol{U}}\right)}{\left\|\mathbf{Res}\left(\overline{\boldsymbol{U}}\right)\right\|_{(V\times Q)'}}, \boldsymbol{V}'\right\rangle_{(V\times Q)',(V\times Q)}.
\end{aligned}$$

Collecting terms with respect to powers of ε gives

$$A_{\boldsymbol{U}}\left(\boldsymbol{U}_1', \boldsymbol{V}'\right) = -\left\langle \frac{\mathbf{Res}\left(\overline{\boldsymbol{U}}\right)}{\left\|\mathbf{Res}\left(\overline{\boldsymbol{U}}\right)\right\|_{(V\times Q)'}}, \boldsymbol{V}'\right\rangle_{(V\times Q)',(V\times Q)}, \tag{8.247}$$

$$A_{\boldsymbol{U}}\left(\boldsymbol{U}_i', \boldsymbol{V}'\right) = -\sum_{j=1}^{i-1} n\left(\boldsymbol{u}_i', \boldsymbol{u}_j', \boldsymbol{v}'\right) \quad i \geq 2.$$

This system is a system of variational problems where the computation of $\boldsymbol{U}_i'$ requires the knowledge of all $\boldsymbol{U}_j'$ with $j < i$. All equations of this system have the same operator on the left-hand side and the coupling occurs only on the right-hand side. □

Remark 8.222 (The Modeling Step) In Bazilevs et al. (2007), it is proposed to truncate the series (8.246) after the first term, i.e.,

$$\boldsymbol{U}' \approx \varepsilon \boldsymbol{U}_1' = \left\|\mathbf{Res}\left(\overline{\boldsymbol{U}}\right)\right\|_{(V\times Q)'} \boldsymbol{U}_1'. \tag{8.248}$$

The function $\boldsymbol{U}_1'$ can be obtained formally by solving the linear partial differential equation (8.247) with the operator $A_{\boldsymbol{U}}\left(\boldsymbol{U}_1', \boldsymbol{V}'\right)$. However, solving (8.247) analytically is in general not possible and the unresolved scale test functions are in practice not available. From the analytical point of view, there is a formal representation of the solution of (8.247) with a so-called fine-scale Green's operator

$$\boldsymbol{U}_1' = G_{\overline{\boldsymbol{U}}}'\left(-\frac{\mathbf{Res}\left(\overline{\boldsymbol{U}}\right)}{\left\|\mathbf{Res}\left(\overline{\boldsymbol{U}}\right)\right\|_{(V\times Q)'}}\right).$$

The proposal in Bazilevs et al. (2007) consists in using a linear approximation of this operator

$$\boldsymbol{U}_1' \approx -\boldsymbol{\delta} \frac{\mathbf{Res}\left(\overline{\boldsymbol{U}}\right)}{\left\|\mathbf{Res}\left(\overline{\boldsymbol{U}}\right)\right\|_{(V\times Q)'}},$$

where $\boldsymbol{\delta}$ is a 4×4 tensor-valued function. Inserting this model in (8.248), the model of the unresolved scales, denoted by $\tilde{\boldsymbol{U}}'$, becomes

$$\tilde{\boldsymbol{U}}' = \varepsilon \tilde{\boldsymbol{U}}_1' = -\boldsymbol{\delta}\mathbf{Res}\left(\overline{\boldsymbol{U}}\right).$$

The choice of $\boldsymbol{\delta}$ will be discussed in Remarks 8.227 and 8.229 as well as in Example 8.228. □

Remark 8.223 (Application to the Navier–Stokes Equations) The approximation of the resolved scales in the two-scale residual-based VMS method, denoted by

$(\boldsymbol{w}^h, r^h)$, is computed in a standard finite element space. It is proposed in Bazilevs et al. (2007) that the parameter $\boldsymbol{\delta}$ is a diagonal tensor-valued functions, i.e.,

$$\boldsymbol{\delta} = \begin{pmatrix} \boldsymbol{\delta}_{\mathrm{m}} & \boldsymbol{0} \\ \boldsymbol{0}^T & \mu \end{pmatrix} = \begin{pmatrix} \delta_{\mathrm{m}} & 0 & 0 & 0 \\ 0 & \delta_{\mathrm{m}} & 0 & 0 \\ 0 & 0 & \delta_{\mathrm{m}} & 0 \\ 0 & 0 & 0 & \mu \end{pmatrix}. \tag{8.249}$$

The model of the unresolved scales has the form

$$\begin{aligned} \tilde{\boldsymbol{U}}' &= -\boldsymbol{\delta}\,\mathbf{Res}\left(\begin{pmatrix} \boldsymbol{w}^h \\ r^h \end{pmatrix}\right) \qquad (8.250) \\ &= -\begin{pmatrix} \boldsymbol{\delta}_{\mathrm{m}}\left(\partial_t \boldsymbol{w}^h - \nu\Delta\boldsymbol{w}^h + \left(\boldsymbol{w}^h\cdot\nabla\right)\boldsymbol{w}^h + \nabla r^h - \boldsymbol{f}\right) \\ \mu\left(\nabla\cdot\boldsymbol{w}^h\right) \end{pmatrix} = -\begin{pmatrix} \mathbf{res}_{\mathrm{m}}^h \\ \mathrm{res}_{\mathrm{c}}^h \end{pmatrix}. \end{aligned}$$

The notation μ for the last component is used because it will turn out that the model of this component leads to a grad-div stabilization term, see Remark 8.226. Now, this model can be inserted in the resolved scale equation (8.238).

In Bazilevs et al. (2007) it is suggested to neglect the models of the terms

$$\left(\partial_t \boldsymbol{u}', \boldsymbol{v}^h\right) \quad \text{and} \quad 2\nu\left(\mathbb{D}\left(\boldsymbol{u}'\right), \mathbb{D}\left(\boldsymbol{v}^h\right)\right).$$

Defining the resolved scales with an appropriate projection, see Remark 8.219, then one of these terms will vanish already in the derivation of the method, the first term if the $L^2(\Omega)$ projection is used and the second term in case of the elliptic projection.

Additionally, the term of the continuity equation with respect to the unresolved scales in (8.238) is integrated by parts

$$\left(\nabla\cdot\boldsymbol{u}', q\right) = -\left(\boldsymbol{u}', \nabla q\right).$$

In this way, there is no derivative of $\boldsymbol{u}'$ in the model.

Inserting (8.250) in (8.238) and using the described modifications gives the resolved scale equation: Find $\boldsymbol{w}^h\ :\ (0,T] \to V^h,\ r^h\ :\ (0,T] \to Q^h$ satisfying

$$\begin{aligned} &\left(\partial_t \boldsymbol{w}^h, \boldsymbol{v}^h\right) + \left(2\nu\mathbb{D}\left(\boldsymbol{w}^h\right), \mathbb{D}\left(\boldsymbol{v}^h\right)\right) + n\left(\boldsymbol{w}^h, \boldsymbol{w}^h, \boldsymbol{v}^h\right) + \left(\nabla\cdot\boldsymbol{w}^h, q^h\right) \\ &- \left(\nabla\cdot\boldsymbol{v}^h, r^h\right) + \left(\mathbf{res}_{\mathrm{m}}^h, \nabla q^h\right) + \left(\mathrm{res}_{\mathrm{c}}^h, \nabla\cdot\boldsymbol{v}^h\right) - n\left(\mathbf{res}_{\mathrm{m}}^h, \boldsymbol{w}^h, \boldsymbol{v}^h\right) \\ &-n\left(\boldsymbol{w}^h, \mathbf{res}_{\mathrm{m}}^h, \boldsymbol{v}^h\right) + n\left(\mathbf{res}_{\mathrm{m}}^h, \mathbf{res}_{\mathrm{m}}^h, \boldsymbol{v}^h\right) = \left\langle \boldsymbol{f}, \boldsymbol{v}^h\right\rangle_{V',V} \end{aligned} \tag{8.251}$$

for all $\left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h$.

The terms $-n\left(\mathbf{res}_{\mathrm{m}}^h, \boldsymbol{w}^h, \boldsymbol{v}^h\right)$ and $-n\left(\boldsymbol{w}^h, \mathbf{res}_{\mathrm{m}}^h, \boldsymbol{v}^h\right)$ can be interpreted as models of the cross terms and $n\left(\mathbf{res}_{\mathrm{m}}^h, \mathbf{res}_{\mathrm{m}}^h, \boldsymbol{v}^h\right)$ is a model for the subgrid scale term, compare Remark 8.129. □

Remark 8.224 (The Trilinear Convective Terms in (8.251)) From the practical point of view, it is advisable that one does not need to compute a derivative of the residual of the momentum equation. For this reason, it is proposed in Bazilevs et al. (2007) to use the following form of the convective term, which is obtained from the divergence form, see Remark 6.6, with integration by parts

$$n(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) = \left(\nabla\cdot\left(\boldsymbol{u}\boldsymbol{v}^T\right),\boldsymbol{w}\right) = -\left(\boldsymbol{u}\boldsymbol{v}^T,\nabla\boldsymbol{w}\right). \tag{8.252}$$

In the convective term of the resolved scales $n\left(\boldsymbol{w}^h,\boldsymbol{w}^h,\boldsymbol{v}^h\right)$ there is no residual and one can use in practice any other form of the convective term described in Remark 6.8.

For the model of the cross terms, one obtains with (8.252), (6.18), and (6.8)

$$\begin{aligned} -\,n\left(\mathbf{res}_{\mathrm{m}}^h,\boldsymbol{w}^h,\boldsymbol{v}^h\right) &= \left(\mathbf{res}_{\mathrm{m}}^h\left(\boldsymbol{w}^h\right)^T,\nabla\boldsymbol{v}^h\right) = \left(\left(\nabla\boldsymbol{v}^h\right)^T\mathbf{res}_{\mathrm{m}}^h,\boldsymbol{w}^h\right)\\ &= \int_\Omega \left(\mathbf{res}_{\mathrm{m}}^h\right)^T\left(\nabla\boldsymbol{v}^h\right)\boldsymbol{w}^h\,d\boldsymbol{x} = \left(\mathbf{res}_{\mathrm{m}}^h,\left(\nabla\boldsymbol{v}^h\right)\boldsymbol{w}^h\right)\\ &= \left(\mathbf{res}_{\mathrm{m}}^h,\left(\boldsymbol{w}^h\cdot\nabla\right)\boldsymbol{v}^h\right) \end{aligned} \tag{8.253}$$

and with (8.252) and (6.18)

$$-\,n\left(\boldsymbol{w}^h,\mathbf{res}_{\mathrm{m}}^h,\boldsymbol{v}^h\right) = \left(\boldsymbol{w}^h\left(\mathbf{res}_{\mathrm{m}}^h\right)^T,\nabla\boldsymbol{v}^h\right) = \left(\mathbf{res}_{\mathrm{m}}^h,\left(\nabla\boldsymbol{v}^h\right)^T\boldsymbol{w}^h\right). \tag{8.254}$$

The model of the subgrid scale term is given by

$$n\left(\mathbf{res}_{\mathrm{m}}^h,\mathbf{res}_{\mathrm{m}}^h,\boldsymbol{v}^h\right) = -\left((\mathbf{res}_{\mathrm{m}}^h\left((\mathbf{res}_{\mathrm{m}}^h\right)^T,\nabla\boldsymbol{v}^h\right). \tag{8.255}$$

□

Remark 8.225 (The SUPG Term in (8.251)) From (8.251) and (8.253), one finds that

$$-n\left(\mathbf{res}_{\mathrm{m}}^h,\boldsymbol{w}^h,\boldsymbol{v}^h\right) + \left(\mathbf{res}_{\mathrm{m}}^h,\nabla q^h\right) = \left(\mathbf{res}_{\mathrm{m}}^h,\left(\boldsymbol{w}^h\cdot\nabla\right)\boldsymbol{v}^h + \nabla q^h\right).$$

This term has just the form of the SUPG term for the convection field $\boldsymbol{w}^h$, see (5.34). Of course, in comparison with a stationary equation, the residual contains the temporal derivative of the velocity. □

Remark 8.226 (The Grad-Div Stabilization in (8.251)) Inserting the concrete formula of the residual of the continuity equation gives the term

$$\left(\mu\nabla\cdot\boldsymbol{w}^h,\nabla\cdot\boldsymbol{v}^h\right),$$

which is just a grad-div stabilization term. □

Remark 8.227 (The Stabilization Parameter $\boldsymbol{\delta}$) As already mentioned, a diagonal tensor is used for $\boldsymbol{\delta}$ with the components δ_{m} and μ, see (8.249). The proposal for choosing δ_{m} and μ in Bazilevs et al. (2007) is based on dimensional arguments and not on numerical analysis. A derivation of the stabilization parameter δ_{m} for compressible flow equations based on such arguments can be found in Shakib et al. (1991). In this paper, a product of a Jacobian matrix, $\boldsymbol{\delta}$, and the transposed of the Jacobian is considered. The dimensional arguments lead to the conclusion that the blocks of this product are dimensionally equivalent to some other matrix. Based on this conclusion, an ansatz for the product is proposed, which contains this matrix, and then the stabilization parameter is derived. Since the whole derivation is somewhat involved, its details will not be presented here but only the results.

Consider parametric finite elements and the bijective map $F_K : \hat{K} \to K$, see Definition B.27. For simplicial meshes, F_K is an affine map of the form (B.18). The inverse map is $F_K^{-1} : K \to \hat{K}$ with $\boldsymbol{x} \mapsto \hat{\boldsymbol{x}}$. Differentiating F_K^{-1} leads to the definition of the symmetric tensor $\mathbb{G}$ with

$$\mathbb{G}_{ij} = \sum_{k=1}^{3} \frac{\partial \hat{x}_k}{\partial x_i} \frac{\partial \hat{x}_k}{\partial x_j}, \quad i,j = 1,2,3.$$

Then, the stabilization parameter proposed in Bazilevs et al. (2007) is given by

$$\delta_{\mathrm{m}} = \left(\frac{4}{\Delta t^2} + \left(\boldsymbol{w}^h\right)^T \mathbb{G} \left(\boldsymbol{w}^h\right) + C_{\mathrm{inv}} \nu^2 \left\|\mathbb{G}\right\|_{\mathrm{F}}^2 \right)^{-1/2}, \tag{8.256}$$

where C_{inv} is the constant in the inverse estimate (C.35).

For the stabilization parameter μ, the vector $\boldsymbol{g}$ with $g_i = \sum_{j=1}^{3} \partial \hat{x}_j / \partial x_i$ is defined and the proposal in Bazilevs et al. (2007) consists in setting

$$\mu = \left(\delta_{\mathrm{m}} \boldsymbol{g}^T \boldsymbol{g}\right)^{-1}. \tag{8.257}$$

The stabilization parameters (8.256) and (8.257) will be discussed in detail for a special case in Example 8.228. □

Example 8.228 (K is a Cube with Edges Parallel to the Axes) Let $\hat{K} = [-1,1]^3$, see Remark B.46, and let K be a cube with edges of length h that are parallel to the coordinate axes. Then the reference map has the form

$$F_K : \hat{K} \to K, \quad \hat{\boldsymbol{x}} \mapsto \frac{1}{2} \begin{pmatrix} h & 0 & 0 \\ 0 & h & 0 \\ 0 & 0 & h \end{pmatrix} \hat{\boldsymbol{x}} + \boldsymbol{b} = \boldsymbol{x}.$$

Considering the inverse map, one finds that

$$\frac{\partial \hat{x}_i}{\partial x_j} = \frac{2}{h} \delta_{ij}, \quad \mathbb{G}_{ij} = \frac{4}{h^2} \delta_{ij}, \quad \left\|\mathbb{G}\right\|_{\mathrm{F}}^2 = \frac{48}{h^4}, \quad \left(\boldsymbol{w}^h\right)^T \mathbb{G} \left(\boldsymbol{w}^h\right) = \frac{4}{h^2} \left\|\boldsymbol{w}^h\right\|_2^2.$$

Then, the stabilization parameter δ_{m} becomes

$$\delta_{\mathrm{m}} = \left(\frac{4}{\Delta t^2} + \frac{4 \left\| \boldsymbol{w}^h \right\|_2^2}{h^2} + \frac{48 C_{\mathrm{inv}} \nu^2}{h^4} \right)^{-1/2}. \tag{8.258}$$

For the grad-div parameter, one obtains $g_i = 2/h$ for $i = 1, 2, 3$, such that $\boldsymbol{g}^T \boldsymbol{g} = 12/h^2$ and

$$\mu = \frac{h^2}{12 \delta_{\mathrm{m}}}. \tag{8.259}$$

Now, the parameters (8.258) and (8.259) will be discussed for the different cases that one of the terms in (8.258) dominates.

- The term $4/\Delta t^2$ dominates in (8.258), i.e., Δt is very small. Then one obtains $\delta_{\mathrm{m}} \sim \Delta t$ and $\mu \sim h^2/\Delta t$.
- The term $4 \left\| \boldsymbol{w}^h \right\|_2^2 / h^2$ dominates in (8.258), i.e., there is a strong convection and $\Delta t \gtrsim h$. In this case, one gets $\delta_{\mathrm{m}} \sim h$ and $\mu \sim h$.
- The term $48 C_{\mathrm{inv}} \nu^2 / h^4$ is dominating in (8.258), i.e., the viscosity dominates or the mesh is very fine and $\Delta t \gtrsim h^2$. This situation leads to $\delta_{\mathrm{m}} \sim h^2$ and $\mu \sim 1$.

Thus, the parameter choice in the second and third case is the same as for equal-order discretizations of the Oseen equations, compare Remark 5.42. In fact, the two-scale residual-based VMS method was applied with such a setting in Bazilevs et al. (2007). □

Remark 8.229 (On the Stabilization Parameters)

- Considering the physical units of the stabilization parameters, one finds that

$$\delta_{\mathrm{m}} \ : \ \left[\left({}^{1}/_{\mathrm{s}^2} + {}^{\mathrm{m}^2}/_{(\mathrm{s}^2 \mathrm{m}^2)} + {}^{\mathrm{m}^4}/_{(\mathrm{s}^2 \mathrm{m}^4)} \right)^{-1/2} \right] = [\mathrm{s}]$$

  and

$$\mu \ \left[\left({}^{\mathrm{s}}/_{\mathrm{m}^2} \right)^{-1} \right] = {}^{\mathrm{m}^2}/_{\mathrm{s}}.$$

  These are the same physical units as in the Oseen equations, see Remark 5.28. Thus, δ_{m} is a time scale and μ is a viscosity scale.
- Concerning the SUPG discretization for time-dependent problems, numerical analysis seems to be available so far only for scalar convection-diffusion equations in John and Novo (2011). For these equations, the analysis for general assumptions on the data leads also to the proposal that the stabilization parameter should scale with the length of the time step, like in (8.256) and (8.258).

- For $\Delta t \to 0$ one finds that $\delta_{\mathrm{m}} \to 0$ and $\mu \to \infty$. A different, heuristic parameter choice which avoids this behavior was discussed in Hsu et al. (2010).
- It can be expected that the parameter in the case of using velocity and pressure finite element spaces that satisfy the discrete inf-sup condition (3.51) has to be chosen in a different way than proposed in Bazilevs et al. (2007). This expectation is based on the different choices for the Oseen equations, see Remark 5.42. In addition, numerical analysis for the transient Oseen equations with grad-div stabilization in de Frutos et al. (2016b) shows that $\mu = \mathcal{O}(1)$ is the asymptotic optimal choice in the convection-dominated regime, in contrast to $\mu = \mathcal{O}(h)$ as it was found in Example 8.228. Since to the best of our knowledge, the two-scale residual-based VMS method was not used so far with inf-sup stable pairs of finite element spaces, the asymptotic correct choice of the stabilization parameter seems to be an open problem in this situation.

□

Remark 8.230 (Time-Dependent Model of the Unresolved Velocity Scales) In Gamnitzer et al. (2010), it is proposed to model the unresolved velocity scales or the subgrid scale velocity with

$$\partial_t \tilde{\boldsymbol{u}}' + \frac{1}{\tilde{\delta}_{\mathrm{m}}} \tilde{\boldsymbol{u}}' = \partial_t \boldsymbol{w}^h - \nu \Delta \boldsymbol{w}^h + \left(\boldsymbol{w}^h \cdot \nabla\right) \boldsymbol{w}^h + \nabla r^h - \boldsymbol{f}, \tag{8.260}$$

instead of (8.250). A time-dependent evolution of the unresolved velocity scales of this form was proposed in Codina (2002) and Codina et al. (2007), see also Sect. 8.8.3 for VMS methods based on time-dependent subgrid scales.

In Gamnitzer et al. (2010), Eq. (8.260) was discretized in a space consisting of bubble functions. The stabilization parameter $\tilde{\delta}_{\mathrm{m}}$ that was proposed in Gamnitzer et al. (2010) possesses the asymptotic $\tilde{\delta}_{\mathrm{m}} = \mathcal{O}(h)$ in the convection-dominated regime. Equal order pairs of finite element spaces, e.g., Q_1/Q_1, were used in the numerical studies in Gamnitzer et al. (2010). These studies were preformed at the turbulent channel flow benchmark problem for $\mathrm{Re}_\tau = 180$ and $\mathrm{Re}_\tau = 395$, see Example D.12. It turned out that in the case of a length of the time step which was not too small, the differences of the results obtained with the steady-state model of the unresolved scales (8.250) and the time-dependent model (8.260) were small. However, for the time-dependent model (8.260), the results were more robust in the sense that the length of the time step did not possess much impact on the results. For the steady-state model, the length of the time step enters the definition of the stabilization parameters (8.256) and (8.257). In particular, δ_{m} becomes small, see Remark 8.229, and a notable impact of the length of the time step on second order statistics was observed. □

Remark 8.231 (Numerical Experience) In Gravemeier et al. (2010), the two-scale residual-based VMS method from Bazilevs et al. (2007) and the algebraic VMS method AVM^3 described in Sect. 8.8.5, both applied with Q_1/Q_1 finite elements, were compared for a turbulent channel flow problem, see Example D.12, and a turbulent flow in a lid driven cavity. With respect to several quantities of interest, the two-scale residual-based VMS method showed less accurate results. In these studies, the simulations with the two-scale residual-based VMS method were also somewhat less efficient. Computational studies in Gravemeier et al. (2011) for a turbulent flow around a cylinder, Example D.13, showed only small differences between the residual-based VMS method and AVM^3. From the point of view of efficency, both VMS methods proved to be clearly superior to the dynamic Smagorinsky model presented in Remark 8.126. □

8.8.3 A Two-Scale VMS Method with Time-Dependent Orthogonal Subscales

Remark 8.232 (Main Idea and Derivation of the Method) The VMS method with orthogonal subscales (OSS-VMS) is based on the abstract two-scale VMS method described in Remark 8.211. It was proposed in Codina (2002).

Let the space for the large scales be the pair of finite element spaces $V^h \times Q^h$. To avoid confusion with the dual spaces, the spaces for the small scales are not marked with a prime but they are denoted by $\tilde{V}' \times \tilde{Q}'$. It is assumed to hold $V = V^h \oplus \tilde{V}'$ and $Q = Q^h \oplus \tilde{Q}'$.

Consider first Eq. (8.238) with the large scale test functions. In the VMS method with orthogonal subscales, the term with the small scales is reformulated applying integration by parts and using (6.24)

$$\begin{aligned}
&A\left(\boldsymbol{u}; \left(\boldsymbol{u}', p'\right), (\overline{\boldsymbol{v}}, \overline{q})\right) \\
&= \left(\partial_t \boldsymbol{u}', \overline{\boldsymbol{v}}\right) + \left(\nu \nabla \boldsymbol{u}', \nabla \overline{\boldsymbol{v}}\right) + n\left(\boldsymbol{u}, \boldsymbol{u}', \overline{\boldsymbol{v}}\right) - \left(\nabla \cdot \overline{\boldsymbol{v}}, p'\right) + \left(\nabla \cdot \boldsymbol{u}', \overline{q}\right) \\
&= \left(\partial_t \boldsymbol{u}', \overline{\boldsymbol{v}}\right) + \left\langle -\nu \Delta \overline{\boldsymbol{v}} - (\boldsymbol{u} \cdot \nabla)\, \overline{\boldsymbol{v}} - \nabla \overline{q}, \boldsymbol{u}'\right\rangle_{V',V} - \left(\nabla \cdot \overline{\boldsymbol{v}}, p'\right) \\
&\quad + \sum_{K \in \mathcal{T}^h} \int_{\partial K} (\nu \nabla \overline{\boldsymbol{v}} + \overline{q} I)\, \boldsymbol{n}_{\partial K} \cdot \boldsymbol{u}'\, d\boldsymbol{s}.
\end{aligned} \tag{8.261}$$

Also for the small scale equation (8.240) integration by parts is applied, leading to the equation

$$\begin{aligned}
&\left\langle \partial_t \boldsymbol{u}' - \nu \Delta \boldsymbol{u}' + \left(\boldsymbol{u}' \cdot \nabla\right) \overline{\boldsymbol{u}} + (\overline{\boldsymbol{u}} \cdot \nabla)\, \boldsymbol{u}' + \left(\boldsymbol{u}' \cdot \nabla\right) \boldsymbol{u}' + \nabla p', \boldsymbol{v}'\right\rangle_{V',V} \\
&\quad + \left(\nabla \cdot \boldsymbol{u}', q'\right) + \sum_{K \in \mathcal{T}^h} \int_{\partial K} \left(\nu \nabla \boldsymbol{u}' - p' I\right) \boldsymbol{n}_{\partial K} \cdot \boldsymbol{v}'\, d\boldsymbol{s} \\
&= -\left\langle \partial_t \overline{\boldsymbol{u}} - \nu \Delta \overline{\boldsymbol{u}} + (\overline{\boldsymbol{u}} \cdot \nabla)\, \overline{\boldsymbol{u}} + \nabla \overline{p} - \boldsymbol{f}, \boldsymbol{v}'\right\rangle_{V',V} - \left(\nabla \cdot \overline{\boldsymbol{u}}, q'\right) \\
&\quad - \sum_{K \in \mathcal{T}^h} \int_{\partial K} (\nu \nabla \overline{\boldsymbol{u}} - \overline{p} I)\, \boldsymbol{n}_{\partial K} \cdot \boldsymbol{v}'\, d\boldsymbol{s}.
\end{aligned} \tag{8.262}$$

The terms with the integrals on the faces sum up to

$$\sum_{K\in\mathcal{T}^h}\int_{\partial K}(\nu\nabla\boldsymbol{u}-pI)\,\boldsymbol{n}_{\partial K}\cdot\boldsymbol{v}'\,d\boldsymbol{s}.$$

It is assumed that the solution $(\boldsymbol{u},p)$ of the continuous problem is sufficiently smooth such that there are no jumps across faces. Then, the term with sum over the faces vanishes.

A first fundamental idea of the method consists in not neglecting the temporal derivative of the small scales but approximating it by

$$\partial_t\boldsymbol{u}'\approx\vartheta\frac{\boldsymbol{u}'-\boldsymbol{u}'_{\mathrm{old}}}{\Delta t},$$

where $\boldsymbol{u}'_{\mathrm{old}}$ is a known approximation of the small scales from a former discrete time and the factor ϑ depends on the used temporal discretization. Additionally, in the trilinear cross term with the subgrid scale velocity as convection, $\boldsymbol{u}'$ is approximated with $\boldsymbol{u}'_{\mathrm{old}}$. Inserting these assumptions and approximations in (8.262) yields

$$\begin{aligned}
&\left\langle\vartheta\frac{\boldsymbol{u}'}{\Delta t}-\nu\Delta\boldsymbol{u}'+(\overline{\boldsymbol{u}}\cdot\nabla)\,\boldsymbol{u}'+\left(\boldsymbol{u}'\cdot\nabla\right)\boldsymbol{u}'+\nabla p',\boldsymbol{v}'\right\rangle_{V',V}+\left(\nabla\cdot\boldsymbol{u}',q'\right)\\
&=-\left\langle-\vartheta\frac{\boldsymbol{u}'_{\mathrm{old}}}{\Delta t}+\partial_t\overline{\boldsymbol{u}}-\nu\Delta\overline{\boldsymbol{u}}+\left(\left(\overline{\boldsymbol{u}}+\boldsymbol{u}'_{\mathrm{old}}\right)\cdot\nabla\right)\overline{\boldsymbol{u}}+\nabla\overline{p}-\boldsymbol{f},\boldsymbol{v}'\right\rangle_{V',V}-\left(\nabla\cdot\overline{\boldsymbol{u}},q'\right)\\
&=-\int_\Omega\mathbf{Res}\left(\overline{\boldsymbol{U}},\boldsymbol{u}'_{\mathrm{old}}\right)\cdot\boldsymbol{V}'\,d\boldsymbol{x}\\
&=-\int_\Omega\left(\mathbf{Res}\left(\overline{\boldsymbol{U}},\boldsymbol{u}'_{\mathrm{old}}\right)+\boldsymbol{V}_{\mathrm{orth}}\right)\cdot\boldsymbol{V}'\,d\boldsymbol{x},
\end{aligned}\tag{8.263}$$

for all $\boldsymbol{V}_{\mathrm{orth}}$ that are $L^2(\Omega)$ orthogonal to $\tilde{V}'\times\tilde{Q}'$.

The ansatz for the small scales has the form, compare (8.242),

$$\boldsymbol{U}'=-\boldsymbol{\delta}\left(\mathbf{Res}\left(\overline{\boldsymbol{U}},\boldsymbol{u}'_{\mathrm{old}}\right)+\boldsymbol{V}_{\mathrm{orth}}\right),\tag{8.264}$$

where $\boldsymbol{\delta}$ is a diagonal tensor of form (8.249).

The second fundamental idea of the method consists in specifying the decomposition into large and small scales in the way that the small scale space should be orthogonal to the large scale space with respect to the inner product of $L^2(\Omega)$. This specification determines $\boldsymbol{V}_{\mathrm{orth}}$, because one obtains with (8.264) for all $\overline{\boldsymbol{V}}\in V^h\times Q^h$

$$0=\left(\boldsymbol{U}',\overline{\boldsymbol{V}}\right)=-\left(\boldsymbol{\delta}\left(\mathbf{Res}\left(\overline{\boldsymbol{U}},\boldsymbol{u}'_{\mathrm{old}}\right)+\boldsymbol{V}_{\mathrm{orth}}\right),\overline{\boldsymbol{V}}\right),\tag{8.265}$$

where this notation assumes that $\mathbf{Res}\left(\overline{\boldsymbol{U}}, \boldsymbol{u}'_{\mathrm{old}}\right)$ is defined in a way such that it belongs to $L^2(\Omega)$. Then, (8.265) defines the $L^2(\Omega)$ projection in the finite element space

$$\boldsymbol{\delta V}_{\mathrm{orth}} = -P^h_{L^2}\left(\boldsymbol{\delta}\mathbf{Res}\left(\overline{\boldsymbol{U}}, \boldsymbol{u}'_{\mathrm{old}}\right)\right)$$

and with (8.264), one obtains

$$\boldsymbol{U}' = -\left(I - P^h_{L^2}\right)\left(\boldsymbol{\delta}\mathbf{Res}\left(\overline{\boldsymbol{U}}, \boldsymbol{u}'_{\mathrm{old}}\right)\right), \tag{8.266}$$

which is the model for the subgrid scales. □

Remark 8.233 (The OSS-VMS Method with Time-Dependent Subscales) To derive a numerical method that can be implemented and used efficiently, a number of simplifications have been proposed in Codina (2002).

- The integrals on the faces of the mesh cells in the large scale Eq. (8.261) are neglected.
- Instead of the dual pairing, a sum over the mesh cells is used in (8.261) and (8.263). The residual defined in this way belongs to $L^2(\Omega)$.
- The Laplacian operators in (8.261) and in the definition of $\mathbf{Res}\left(\overline{\boldsymbol{U}}, \boldsymbol{u}'_{\mathrm{old}}\right)$ in (8.263) are neglected.
- Instead of the projection (8.266), the small scales are defined by the standard $L^2(\Omega)$ projection

$$\boldsymbol{U}'|_K = -\boldsymbol{\delta}_K\left[\left(I - P^h_{L^2}\right)\left(\mathbf{Res}\left(\overline{\boldsymbol{U}}, \boldsymbol{u}'_{\mathrm{old}}\right)\right)\right]\Big|_K.$$

If $\boldsymbol{\delta}_K = \boldsymbol{\delta}$ is constant on Ω, this expression is the same as (8.266).

Note that because of the $L^2(\Omega)$ orthogonality of the subgrid scales and the large scales, the term $(\partial_t \boldsymbol{u}', \overline{\boldsymbol{v}})$ in (8.261) and the term $(\partial_t \overline{\boldsymbol{u}}, \boldsymbol{v}')$ in (8.263) vanish. For the same reason, it is $\left(I - P^h_{L^2}\right)\boldsymbol{u}'_{\mathrm{old}} = \boldsymbol{u}'_{\mathrm{old}}$.

Using the definition (8.263) of the residual, reordering terms, and using the notation $(\boldsymbol{w}^h, r^h)$ for the approximation of the large scales gives the model

$$\begin{aligned}
&\left(\partial_t \boldsymbol{w}^h, \boldsymbol{v}^h\right) + \nu\left(\nabla \boldsymbol{w}^h, \nabla \boldsymbol{v}^h\right) + n\left(\boldsymbol{w}^h + \boldsymbol{w}', \boldsymbol{w}^h, \boldsymbol{v}^h\right) - \left(\nabla\cdot\boldsymbol{v}^h, r^h\right) + \left(\nabla\cdot\boldsymbol{w}^h, r^h\right)\\
&\quad + \sum_{K\in\mathcal{T}^h}\Bigg(\left(I - P^h_{L^2}\right)\left(\left(\left(\boldsymbol{w}^h + \boldsymbol{w}'\right)\cdot\nabla\right)\boldsymbol{w}^h + \nabla r^h\right),\\
&\qquad\qquad \delta_{\mathrm{m},K}\left(\left(\left(\boldsymbol{w}^h + \boldsymbol{w}'\right)\cdot\nabla\right)\boldsymbol{v}^h + \nabla q^h\right)\Bigg)_K\\
&\quad + \sum_{K\in\mathcal{T}^h}\left(\left(I - P^h_{L^2}\right)\left(\nabla\cdot\boldsymbol{w}^h\right), \mu_K \nabla\cdot\boldsymbol{v}^h\right)_K
\end{aligned} \tag{8.267}$$

$$= \left(\boldsymbol{f}, \boldsymbol{v}^h\right) + \sum_{K\in\mathcal{T}^h} \left((I - P_{L^2}^h)\boldsymbol{f}, \delta_{\mathrm{m},K} \left(\left(\left(\boldsymbol{w}^h + \boldsymbol{w}'\right)\cdot\nabla\right)\boldsymbol{v}^h + \nabla q^h\right)\right)_K$$
$$+ \frac{\vartheta}{\Delta t} \sum_{K\in\mathcal{T}^h} \left(\boldsymbol{w}'_{\mathrm{old}}, \delta_{\mathrm{m},K} \left(\left(\left(\boldsymbol{w}^h + \boldsymbol{w}'\right)\cdot\nabla\right)\boldsymbol{v}^h + \nabla q^h\right)\right)_K .$$

Model (8.267) needs $\boldsymbol{w}'$ explicitly, which is given by

$$\boldsymbol{w}'|_K = \delta_{\mathrm{m},K} \left(\vartheta \frac{\boldsymbol{w}'_{\mathrm{old}}}{\Delta t} - \left(I - P_{L^2}^h\right)\left(\left(\left(\boldsymbol{w}^h + \boldsymbol{w}'\right)\cdot\nabla\right)\boldsymbol{w}^h + \nabla r^h - \boldsymbol{f}\right)\right)\Bigg|_K . \tag{8.268}$$

To keep the subgrid scales in the convection field of the convective terms can be considered as the third main idea of the method. The convection field $\left(\boldsymbol{w}^h + \boldsymbol{w}'\right)$ is called advection velocity. □

Remark 8.234 (On the OSS-VMS Method with Time-Dependent Subscales)

- An implementation of this method with concrete time stepping schemes and using a Picard iteration is described in Codina (2002).
- The sum over the mesh cells on the left-hand side of (8.267) includes a pressure-pressure coupling, e.g., as in the PSPG method presented in Sect. 4.5.1. Thus, the OSS-VMS method can be used with pairs of finite element spaces that do not satisfy the discrete inf-sup condition (3.51). In fact, usually such pairs of spaces were applied in the numerical studies that can be found in the literature.
- The cross terms and the subgrid scale term of the sgs tensor, compare (8.29) and Remark 8.129, are modeled in the OSS-VMS method. None of them is neglected.
- The subscales have to be stored. They are needed to assemble matrix entries and right-hand side entries in the OSS-VMS method (8.267). Thus, it suffices to store these scales in the quadrature points. The update (8.268) can be computed in each quadrature point.
- Proposals for stabilization parameters $\delta_{\mathrm{m},K}$ and μ_K were derived in Codina (2002) on the basis of a Fourier analysis:

$$\delta_{\mathrm{m},K} = \left\{\frac{1}{\vartheta\Delta t} + \left[\left(C_1\frac{\nu}{h^2}\right)^2 + \left(C_2\frac{\left\|\boldsymbol{w}^h + \boldsymbol{w}'\right\|_2}{h}\right)^2\right]^{1/2}\right\}^{-1},$$
$$\mu_K = \left[\nu^2 + \left(\frac{C_2}{C_1}\frac{\left\|\boldsymbol{w}^h + \boldsymbol{w}'\right\|_2}{h}\right)^2\right]^{1/2},$$

  with user-chosen constants $C_1, C_2 > 0$. Considering the convection-dominated regime and time steps that are not very small, one finds that $\delta_{\mathrm{m},K} \sim h$ and $\mu_K \sim h$, which corresponds to the asymptotic optimal choice of the stabilization parameters in the SUPG/PSPG/grad-div method for the Oseen equations and equal-order pairs, see Remark 5.42. Numerical studies investigating different choices of C_1, C_2 were performed in Colomés et al. (2015).

- Note that the OSS-VMS method applies a global $L^2(\Omega)$ projection and not a local projection as LPS methods, see Remark 5.52.
- In Codina (2002), other algorithmic aspects are discussed, e.g., to compute the subscales with an explicit time stepping scheme or to neglect the temporal evolution of the subscales at all, i.e., setting $\boldsymbol{w}'_{\text{old}} = \mathbf{0}$ in (8.267), (8.268). The latter approach is called static or quasi-static subscales.
- There exist modifications and extensions of the prototype OSS-VMS method (8.267), (8.268), e.g., see Colomés et al. (2015).

□

Remark 8.235 (Energy Balance) The energy balance of the OSS-VMS method was studied in Principe et al. (2010) and Codina et al. (2011). Assuming that a solution exists and using this solution as test function, it turns out that there is also a scale separation in the local kinetic energy, i.e., there are separate balances for the large scales and the subscales because the temporal derivative of one kind of scales does not enter the energy balance of the other kind of scales. This result holds only if the subscales are $L^2(\Omega)$ orthogonal to the finite element space since in this case the respective terms vanish, see Remark 8.233. □

Remark 8.236 (Backscatter of Energy) The term

$$\left\langle -\nu\Delta\overline{\boldsymbol{v}} - (\boldsymbol{u}\cdot\nabla)\,\overline{\boldsymbol{v}} - \nabla\overline{q}, \boldsymbol{u}'\right\rangle_{V',V} \tag{8.269}$$

in (8.261) can be considered to describe the energy exchange between the large scales and the subscales. For simplicity, $\boldsymbol{f}\in V^h$ is assumed, such that $\left(I-P_{L^2}^h\right)\boldsymbol{f} = \mathbf{0}$, $\delta_{\text{m},K} = \delta_{\text{m}}$ is considered, the Laplacian in (8.269) is neglected, and the dual pairing is replaced by a sum over the mesh cells. Setting $\left(\boldsymbol{v}^h, q^h\right) = \left(\boldsymbol{w}^h, r^h\right)$ in (8.268), inserting the subscales (8.268), using $\boldsymbol{w} = \boldsymbol{w}^h + \boldsymbol{w}'$, and using the $L^2(\Omega)$ orthogonality of the different scales gives

$$\begin{aligned}
&\delta_{\text{m}} \sum_{K\in\mathcal{T}^h} \left(-(\boldsymbol{w}\cdot\nabla)\,\boldsymbol{w}^h - \nabla r^h, -\left(I-P_{L^2}^h\right)\left(\left(\boldsymbol{w}^h+\boldsymbol{w}'\right)\cdot\nabla\right)\boldsymbol{w}^h + \nabla r^h\right)_K \\
&\qquad\qquad + \delta_{\text{m}} \sum_{K\in\mathcal{T}^h} \left(-(\boldsymbol{w}\cdot\nabla)\,\boldsymbol{w}^h - \nabla r^h, \vartheta\frac{\boldsymbol{w}'_{\text{old}}}{\Delta t}\right)_K \\
&\qquad = \delta_{\text{m}} \left\|\left(I-P_{L^2}^h\right)\left(\left(\boldsymbol{w}^h+\boldsymbol{w}'\right)\cdot\nabla\right)\boldsymbol{w}^h\right\|_{L^2(\Omega)}^2 \\
&\qquad\qquad - \delta_{\text{m}} \left(\left(I-P_{L^2}^h\right)\left((\boldsymbol{w}\cdot\nabla)\,\boldsymbol{w}^h + \nabla r^h\right), \vartheta\frac{\boldsymbol{w}'_{\text{old}}}{\Delta t}\right).
\end{aligned} \tag{8.270}$$

Backscatter of energy is given if (8.270) is negative. The first term on the right-hand side of (8.270) is non-negative. It follows that backscatter can only occur if the subscales are time-dependent. Otherwise, the second term on the right-hand side of (8.270) vanishes since $\boldsymbol{w}'_{\text{old}} = \boldsymbol{0}$. For time-dependent subscales, the sign of the this term is not clear. In numerical studies in Principe et al. (2010), it was observed that in fact backscatter occurred locally in space and time using the OSS-VMS method. □

Remark 8.237 (Finite Element Error Analysis) A finite element error analysis for a stabilized method with orthogonal subscales is presented in Codina (2008). In this paper, the Oseen Eq. (5.1), with $c = 0$, are considered and the analysis is performed for equal-order pairs of finite element spaces. In contrast to (8.267), (8.268), not the standard $L^2(\Omega)$ projection is used in the definition of the method but the weighted $L^2(\Omega)$ projection with the stabilization parameters inside, see (8.266). As usual for stabilized methods, the analysis is performed for a norm that involves contributions from the stabilization terms. It turns out that the bilinear form of the method is not coercive in this norm. Thus, for obtaining the existence and uniqueness of a solution, an inf-sup condition is proved, see Lemma 5.38 and Corollary 5.40 where this approach is applied for the SUPG/PSPG/grad-div method. Finally, it is shown that the method converges with optimal order if the stabilization parameters are $\delta_{\text{m},K} = \mathcal{O}(h_K)$, $\mu_K = \mathcal{O}(h_K)$ and if the mesh width is sufficiently small.

Similarly as for the SUPG/PSPG/grad-div method, the norm that was used in the error analysis possesses a term with the sum of the streamline derivative of the velocity and the gradient of the pressure, see (5.40) and Theorem 5.41. There is no separate control of these two terms. For this reason, in Codina (2008) a second method is considered whose definition uses individual projections of the streamline derivative of the velocity and the gradient of the pressure. For this method, it is possible to control both terms separately and to show the optimal order of convergence.

The long time behavior of the OSS-VMS method with dynamic subscales is investigated in Badia et al. (2010). As a first step, the existence and uniqueness of a solution of the considered continuous-in-time version of the method is proved. Then, the stability of the solution is studied. With standard assumptions, the stability of the finite element solution and also of the subscales is shown for a bounded time interval. Using slightly stronger assumptions, also the stability for $T \to \infty$ is proved, e.g., giving $\boldsymbol{w}^h \in L^\infty\left(0, \infty; L^2(\Omega)\right)$ and $\boldsymbol{w}' \in L^\infty\left(0, \infty; L^2(\Omega)\right)$. These results imply the existence of an absorbing set in $L^2(\Omega)$ for both $\boldsymbol{w}^h$ and $\boldsymbol{w}'$. In the two-dimensional case, also the existence of an absorbing set for $\boldsymbol{w}^h$ in $H^1(\Omega)$ is proved in Badia et al. (2010). From this result, one can conclude that, for $\boldsymbol{f}$ being time-independent, the absorbing set of $\boldsymbol{w}^h$ in $L^2(\Omega)$ is even a global attractor for $d = 2$. For the analysis presented in Badia et al. (2010) it was crucial that the subscales are time-dependent. □

Remark 8.238 (The Algebraic Subgrid Scale (ASGS) VMS Method) The ASGS-VMS method, whose prototype was derived in the framework of a two-scale VMS method in Codina (2001b), see also Codina et al. (2007), can be defined by choosing

$\boldsymbol{V}_{\text{orth}} = \mathbf{0}$ in (8.264), instead of employing the orthogonality of the subgrid scale space and the finite element space for defining $\boldsymbol{V}_{\text{orth}}$. Since these spaces are not orthogonal, the terms $(\partial_t \boldsymbol{u}', \overline{\boldsymbol{v}})$ and $(\partial_t \overline{\boldsymbol{u}}, \boldsymbol{v}')$ do not vanish. With similar assumptions as for the OSS-VMS method, one obtains the prototype ASGS-VMS method

$$\begin{aligned}
&\left(\partial_t \boldsymbol{w}^h, \boldsymbol{v}^h\right) + \nu\left(\nabla \boldsymbol{w}^h, \nabla \boldsymbol{v}^h\right) + n\left(\boldsymbol{w}^h + \boldsymbol{w}', \boldsymbol{w}^h, \boldsymbol{v}^h\right) - \left(\nabla\cdot \boldsymbol{v}^h, r^h\right) \\
&\quad + \left(\nabla\cdot \boldsymbol{w}^h, r^h\right) + \left(\partial_t \boldsymbol{w}', \boldsymbol{v}^h\right) \\
&\quad + \sum_{K\in\mathcal{T}^h} \left(\left(\left(\boldsymbol{w}^h + \boldsymbol{w}'\right)\cdot\nabla\right)\boldsymbol{w}^h + \nabla r^h, \delta_{\text{m},K}\left(\left(\left(\boldsymbol{w}^h + \boldsymbol{w}'\right)\cdot\nabla\right)\boldsymbol{v}^h + \nabla q^h\right)\right)_K \\
&\quad + \sum_{K\in\mathcal{T}^h} \left(\nabla\cdot \boldsymbol{w}^h, \mu_K \nabla\cdot \boldsymbol{v}^h\right)_K \\
&= \left(\boldsymbol{f}, \boldsymbol{v}^h\right) + \sum_{K\in\mathcal{T}^h} \left(\boldsymbol{f}, \delta_{\text{m},K}\left(\left(\left(\boldsymbol{w}^h + \boldsymbol{w}'\right)\cdot\nabla\right)\boldsymbol{v}^h + \nabla q^h\right)\right)_K \\
&\quad + \frac{\vartheta}{\Delta t}\sum_{K\in\mathcal{T}^h} \left(\boldsymbol{w}'_{\text{old}}, \delta_{\text{m},K}\left(\left(\left(\boldsymbol{w}^h + \boldsymbol{w}'\right)\cdot\nabla\right)\boldsymbol{v}^h + \nabla q^h\right)\right)_K
\end{aligned}$$

with

$$\boldsymbol{w}'|_K = \delta_{\text{m},K}\left(\vartheta \frac{\boldsymbol{w}'_{\text{old}}}{\Delta t} - \left(\partial_t \boldsymbol{w}^h + \left(\boldsymbol{w}^h + \boldsymbol{w}'\right)\cdot\nabla\right)\boldsymbol{w}^h + \nabla r^h - \boldsymbol{f}\right)\Bigg|_K .$$

Refinements of this method, taking into account for instance the discretizations of temporal derivatives in the residual, can be found, e.g., in Colomés et al. (2015).

Note that the principles for deriving the quasi-static version of the ASGS-VMS method, i.e., with $\boldsymbol{w}'_{\text{old}} = \mathbf{0}$, are more or less the same as for deriving the residual-based VMS method presented in Sect. 8.8.2, such that both methods are similar. □

Remark 8.239 (Numerical Experience) There are a number of comparisons of the OSS-VMS method and the ASGS-VMS method. All of them were performed for equal-order pairs of finite element spaces. The most detailed assessment of these methods can be found in Colomés et al. (2015), where for instance turbulent channel flow problems, see Example D.12, were studied. By construction, it is to be expected that the OSS-VMS method introduces less numerical viscosity than the ASGS-VMS method, since in the former method only the fluctuation (identity operator minus projection operator) of the residual occurs in the stabilization terms. This expectation is often met in the numerical results, e.g., see Codina (2008).

In Colomés et al. (2015), it was found that overall the OSS-VMS and ASGS-VMS methods give similar results. Both methods converged to reference solutions when the mesh was refined or the polynomial degree of the finite element spaces was increased. Furthermore, it was observed that both methods are sensitive to the concrete choice of the stabilization parameters $\delta_{\text{m},K}$ and μ_K. In fact, this choice is more important for the numerical results than the choice of the VMS method itself. With respect to the computational cost, the OSS-VMS method with dynamic subscales proved to be most efficient. □

8.8.4 A Three-Scale Bubble VMS Method

Remark 8.240 (Realizations of Three-Scale Bubble VMS Methods) Bubble VMS methods can be considered as the most direct realization of a three-scale VMS method as described in Remark 8.214.

The main goal of using bubble functions for approximating the small resolved scales consists in splitting the Eq. (8.245) for these scales into a number of local problems to obtain an efficient method, see Remark 8.242. This idea was already pointed out in Hughes et al. (2000).

A first realization of this idea can be found in Gravemeier et al. (2004, 2005) and Gravemeier (2006c). In these papers, the velocity and the pressure were approximated with bilinear or trilinear finite elements. Only the velocity space was enriched with bubble functions for the small resolved scales. With this enrichment, the pair of finite element spaces becomes inf-sup stable. The stabilizing effect of bubble functions with respect to the discrete inf-sup condition (3.51) was already seen for the MINI element in Sect. 3.6.1 or the pair $P_2^{\text{bubble}}/P_1^{\text{disc}}$, see Remark 3.133. The model for the small resolved pressure does not use bubble functions, compare Remark 8.244. A main issue in Gravemeier et al. (2004, 2005) and Gravemeier (2006c) was the investigation of the turbulence model applied to the small resolved scales. A realization of a three-scale bubble VMS method with second order velocity and first order pressure, which followed the principal ideas of Gravemeier et al. (2004, 2005), was explored in John and Kindl (2010a).

In this section, just one possible approach for a bubble VMS method is sketched. □

Remark 8.241 (Bubble VMS Method: Basic Equations) Let the resolved scales $(\boldsymbol{u}^h, p^h)$ be decomposed into large scales $(\overline{\boldsymbol{u}}, \overline{p})$ and small resolved scales $(\hat{\boldsymbol{u}}, \hat{p})$.

The equation with the large scale test functions, where the coupling of the large scales and the unresolved scales has been neglected, has the form, compare (8.244),

$$\left(\partial_t \boldsymbol{u}^h, \overline{\boldsymbol{v}}\right) + \left(2\nu \mathbb{D}\left(\boldsymbol{u}^h\right), \mathbb{D}\left(\overline{\boldsymbol{v}}\right)\right) + n\left(\boldsymbol{u}^h, \boldsymbol{u}^h, \overline{\boldsymbol{v}}\right) - \left(\nabla \cdot \overline{\boldsymbol{v}}, p^h\right) + \left(\nabla \cdot \boldsymbol{u}^h, \overline{q}\right) = \left(\boldsymbol{f}, \overline{\boldsymbol{v}}\right).$$

Applying the splitting of the resolved scales yields

$$\begin{aligned} &(\partial_t \overline{\boldsymbol{u}}, \overline{\boldsymbol{v}}) + (2\nu \mathbb{D}(\overline{\boldsymbol{u}}), \mathbb{D}(\overline{\boldsymbol{v}})) + n(\overline{\boldsymbol{u}}, \overline{\boldsymbol{u}}, \overline{\boldsymbol{v}}) - (\nabla \cdot \overline{\boldsymbol{v}}, \overline{p}) + (\nabla \cdot \overline{\boldsymbol{u}}, \overline{q}) \\ &= (\boldsymbol{f}, \overline{\boldsymbol{v}}) - \Big[(\partial_t \hat{\boldsymbol{u}}, \overline{\boldsymbol{v}}) + (2\nu \mathbb{D}(\hat{\boldsymbol{u}}), \mathbb{D}(\overline{\boldsymbol{v}})) \\ &\quad + n\left(\boldsymbol{u}^h, \hat{\boldsymbol{u}}, \overline{\boldsymbol{v}}\right) + n(\hat{\boldsymbol{u}}, \overline{\boldsymbol{u}}, \overline{\boldsymbol{v}}) - (\nabla \cdot \overline{\boldsymbol{v}}, \hat{p}) + (\nabla \cdot \hat{\boldsymbol{u}}, \overline{q}) \Big]. \end{aligned} \tag{8.271}$$

Similarly, one obtains an equation for the small resolved scale test functions, where the eddy viscosity model is already included,

$$\begin{aligned}
&(\partial_t \hat{\boldsymbol{u}}, \hat{\boldsymbol{v}}) + ((2\nu + \nu_{\mathrm{T}}) \mathbb{D}(\hat{\boldsymbol{u}}), \mathbb{D}(\hat{\boldsymbol{v}})) \\
&\quad + n\left(\boldsymbol{u}^h, \hat{\boldsymbol{u}}, \hat{\boldsymbol{v}}\right) - (\nabla \cdot \hat{\boldsymbol{v}}, \hat{p}) + (\nabla \cdot \hat{\boldsymbol{u}}, \hat{q}) \\
&= (\boldsymbol{f}, \hat{\boldsymbol{v}}) - \Big[(\partial_t \overline{\boldsymbol{u}}, \hat{\boldsymbol{v}}) + (2\nu \mathbb{D}(\overline{\boldsymbol{u}}), \mathbb{D}(\hat{\boldsymbol{v}})) \\
&\quad + n\left(\boldsymbol{u}^h, \overline{\boldsymbol{u}}, \hat{\boldsymbol{v}}\right) - (\nabla \cdot \hat{\boldsymbol{v}}, \overline{p}) + (\nabla \cdot \overline{\boldsymbol{u}}, \hat{q}) \Big].
\end{aligned} \tag{8.272}$$

□

Remark 8.242 (Localization of the Small Resolved Velocity Scale Equation with Bubble Functions) In a bubble-based finite element VMS method, standard finite element spaces are used for the large scales $\overline{V} \times \overline{Q} = V^h \times Q^h$. The finite element spaces for the small resolved scales require a higher resolution than the finite element spaces for the large scales. This goal can be achieved in various ways: by using higher order finite elements, by refining the given grid, or by combining these approaches. However, the result of all approaches is that the solution of the small resolved scale Eq. (8.272) would be much more expensive than solving the large scale Eq. (8.271). This difficulty is circumvented in a bubble-based finite element VMS method by considering (8.272) in a space of bubble functions for the velocity. A bubble function is a function from $H_0^1(\Omega)$ whose support is only one mesh cell and which vanishes on the faces of this mesh cell. With these functions, the solution of (8.272) can be localized. Because of the modeling of the small resolved pressure explained in Remark 8.244, only a bubble space for the small resolved velocity is needed. In practice, this space has to be finite-dimensional and it will be denoted by $\widehat{V}^h_{\mathrm{bub}}$. □

Remark 8.243 (An Unphysical Property Introduced by Using Bubble Functions for Modeling the Small Resolved Scales) There is a principal question in all bubble-based VMS approaches concerning the physics of the modeling of the small resolved scales. Since these scales are represented by bubble functions, they can move within a mesh cell but they cannot move directly from one mesh cell to another because of the homogeneous Dirichlet boundary conditions on the faces of the mesh cells. The information contained in the small resolved scales can be distributed to other mesh cells only indirectly by the coupling of the small resolved scales to the large scales. This quasi-stationary modeling of the small resolved scales does not reflect the physical reality. However, there are no numerical studies available that investigate the impact of this unphysical modeling in detail. □

Remark 8.244 (Modeling of the Small Resolved Pressure) It was proposed in Gravemeier et al. (2004, 2005) to model the small resolved scale pressure in the form

$$\hat{p} = - \sum_{K \in \mathcal{T}^h} \mu_K (\nabla \cdot \overline{\boldsymbol{u}}). \tag{8.273}$$

In this way, the influence of the small resolved scale pressure onto the large scales is not directly taken into account but this influence is modeled. Using (8.273), $\hat{p}$ vanishes from the small resolved scale Eq. (8.272). Since the small resolved pressure disappeared, the imposition of a divergence constraint for the small resolved velocity is not possible. Since there is no longer a divergence constraint for $\hat{\boldsymbol{u}}$, it does not make sense to have a term with this function in the divergence constraint of the large scale Eq. (8.271). Altogether, all terms in (8.271) and (8.272) coming from the divergence constraint that include small resolved scales will be neglected by setting

$$(\nabla \cdot \hat{\boldsymbol{u}}, \overline{q}) = (\nabla \cdot \overline{\boldsymbol{u}}, \hat{q}) = (\nabla \cdot \hat{\boldsymbol{u}}, \hat{q}) = 0. \tag{8.274}$$

One obtains from (8.272) a vector-valued equation for $\hat{\boldsymbol{u}}$. The model (8.273) for the small resolved scale pressure is included in the large scale equation leading to a grad-div stabilization term.

A model of form (8.273) was also used in the two-scale residual-based VMS method, see (8.250). It has the interpretation that the small (resolved) pressure is driven by the residual of the large scale continuity equation. □

Remark 8.245 (Further Simplifications) Usually, some additional simplifying assumptions are made for the terms with the small resolved velocity scales. The equation for the small resolved velocity scales is only solved once in each discrete time, at the beginning, giving the solution $\hat{\boldsymbol{u}}^{(1)}$. Consequently, this equation is linearized and all terms with $\overline{\boldsymbol{u}}$ are treated explicitly. For reasons of efficiency, the gradient form of the viscous term is used in the small resolved scale equation and some right-hand side terms in the large scale equation. In particular, the small resolved scale equation decouples into three scalar equations since the system matrix becomes a block diagonal matrix, see Remarks 4.66 and 4.69.

Because the equation for the small resolved scales is solved only at the beginning of each time step, the temporal derivatives in (8.271) and (8.272) have to be modified. For the large scale Eq. (8.271), one uses

$$\partial_t \hat{\boldsymbol{u}} \approx \frac{\hat{\boldsymbol{u}}_{n+1} - \hat{\boldsymbol{u}}_n}{\Delta t_{n+1}} \approx \frac{\hat{\boldsymbol{u}}^{(1)} - \hat{\boldsymbol{u}}_n}{\Delta t_{n+1}}.$$

In the small resolved scale equations, one assumes that the temporal change in the large scales can be neglected, i.e., that $\partial_t \overline{\boldsymbol{u}} = 0$. □

Remark 8.246 (Bubble VMS Method with Time-Dependent Small Resolved Velocity Scales) Inserting the models and simplifications from Remarks 8.244 and 8.245 in (8.271) and (8.272) and using the convention for the notation from Remark 8.217 leads to the following system of equations: Find $\overline{\boldsymbol{w}} : (0, T] \to \overline{V}$, $\overline{r} : (0, T] \to \overline{Q}$ satisfying

$$\begin{aligned}&(\partial_t \overline{\boldsymbol{w}}, \overline{\boldsymbol{v}}) + \left(2\nu \mathbb{D}\left(\overline{\boldsymbol{w}}\right), \mathbb{D}\left(\overline{\boldsymbol{v}}\right)\right)\\ &\quad + n\left(\overline{\boldsymbol{w}}, \overline{\boldsymbol{w}}, \overline{\boldsymbol{v}}\right) - (\nabla \cdot \overline{\boldsymbol{v}}, \overline{r}) + (\nabla \cdot \overline{\boldsymbol{w}}, \overline{q}) + \sum_{K \in \mathcal{T}^h} \mu_K (\nabla \cdot \overline{\boldsymbol{w}}, \nabla \cdot \overline{\boldsymbol{v}})_K\end{aligned}$$

$$= (\boldsymbol{f}, \overline{\boldsymbol{v}}) - \Bigg[\left(\frac{\hat{\boldsymbol{w}}^{(1)} - \hat{\boldsymbol{w}}_n}{\Delta t_{n+1}}, \overline{\boldsymbol{v}} \right) + \left(\nu \nabla \hat{\boldsymbol{w}}^{(1)}, \nabla \overline{\boldsymbol{v}} \right) \tag{8.275}$$

$$n\left(\overline{\boldsymbol{w}}_n + \hat{\boldsymbol{w}}^{(1)}, \hat{\boldsymbol{w}}^{(1)}, \overline{\boldsymbol{v}} \right) + n\left(\hat{\boldsymbol{w}}^{(1)}, \overline{\boldsymbol{w}}_n, \overline{\boldsymbol{v}} \right) \Bigg]$$

for all $(\overline{\boldsymbol{v}}, \overline{p}) \in \overline{V} \times \overline{Q}$. The equation for computing $\hat{\boldsymbol{w}}^{(1)} : (0, T] \to \widehat{V}^h_{\text{bub}}$ reads as follows

$$\begin{aligned} &\left(\partial_t \widehat{\boldsymbol{w}}^{(1)}, \widehat{\boldsymbol{v}} \right) + \left((\nu + \nu_{\text{T}}) \nabla \widehat{\boldsymbol{w}}^{(1)}, \nabla \widehat{\boldsymbol{v}} \right) + n\left(\boldsymbol{w}_n^h, \widehat{\boldsymbol{w}}^{(1)}, \widehat{\boldsymbol{v}} \right) \\ &= (\boldsymbol{f}, \widehat{\boldsymbol{v}}) - \Bigg[(\nu \nabla \overline{\boldsymbol{w}}_n, \nabla \widehat{\boldsymbol{v}}) + n\left(\boldsymbol{w}_n^h, \overline{\boldsymbol{w}}_n, \widehat{\boldsymbol{v}} \right) \\ &\quad - (\nabla \cdot \widehat{\boldsymbol{v}}, \overline{r}_n) + \sum_{K \in \mathcal{T}^h} \mu_K (\nabla \cdot \overline{\boldsymbol{w}}_n, \nabla \cdot \widehat{\boldsymbol{v}})_K \Bigg] \end{aligned} \tag{8.276}$$

for all $\hat{\boldsymbol{v}} \in \widehat{V}^h_{\text{bub}}$. The subscript n refers always to functions computed in the previous discrete time.

Equation (8.276) can be interpreted in the way that the small resolved velocity scales are driven by the residual of the large scale momentum equation. □

Remark 8.247 (Other Bubble VMS Methods) The way for defining a bubble VMS method described in this section is just one possible approach. Other simplifications are possible, leading to (slightly) different equations compared with (8.275) and (8.276). For instance, in John and Kindl (2010a) a bubble VMS method was studied with quasi-stationary small resolved scales that avoids the storage of the bubble velocity from the previous discrete time, as it is necessary in (8.275), (8.276). □

Remark 8.248 (Residual-Free Bubble Methods) The use of bubble functions for stabilizing dominant convection was already proposed prior to VMS methods. Exactly as described in Remarks 8.244 and 8.246, these bubble functions solve equations with the residual obtained with some standard finite element method. For this reason, this approach is called residual-free bubble (RFB) method. This idea was first proposed for scalar convection-diffusion equations in Brezzi and Russo (1994) and applications to laminar incompressible flows can be found, e.g., in Franca and Nesliturk (2001). Thus, the bubble VMS method can be considered as a generalization of the RFB method in the sense that a turbulence model is introduced in the equation for the small resolved velocity scales to account for the turbulent character of the flow. □

Remark 8.249 (Choice of the Turbulence Model) The definition of the small resolved scale Eq. (8.276) requires the choice of the turbulence model ν_{T}. In Gravemeier et al. (2004, 2005), a dynamic Smagorinsky model, see Remark 8.126, was used. The studies in John and Kindl (2010a) applied static Smagorinsky models of the form

$$\nu_{\mathrm{T}} = C_S h_K^2 \left\| \mathbb{D}\left(\overline{\boldsymbol{w}} + \hat{\boldsymbol{w}}^{(1)}\right)\right\|_{\mathrm{F}} \quad \text{and} \quad \nu_{\mathrm{T}} = C_S h_K^2 \left\| \mathbb{D}\left(\overline{\boldsymbol{w}}\right)\right\|_{\mathrm{F}}. \tag{8.277}$$

□

Remark 8.250 (Approximating the Small Resolved Velocity Scales with Bubble Functions) The space $\widehat{V}^h_{\mathrm{bub}}$ has to be specified in order to compute the solution of the bubble Eq. (8.276). In Gravemeier et al. (2004, 2005), Gravemeier (2006c) and John and Kindl (2010a), local grids in each hexahedral mesh cell were used. A typical size of a local grid was $5 \times 5 \times 5$ sub cells. On these local grids, the equation for the bubble functions was discretized, usually with Q_1 finite elements. For the use of the dynamic Smagorinsky model in Gravemeier et al. (2004, 2005), a second local grid was applied that was somewhat finer than the first local grid. □

Remark 8.251 (Numerical Experience with Bubble VMS Methods) In John and Kindl (2010a), it is mentioned that the application of a bubble finite element VMS method is quite complicated: one has to decide about the simplifying assumptions with respect to the small resolved scales and also the implementation is quite involved. In addition, it turned out that the dominating term of the model is the grad-div term that evolves from modeling the small resolved pressure, see (8.274). Using only this term without modeling the small resolved velocity led to stable simulations. However, applying in addition to the grad-div stabilization also the bubble-based model for the small resolved velocity improved the accuracy of the results. It is also mentioned in John and Kindl (2010a) that the rather coarse grids for solving the localized Eq. (8.276) required to take large values for the parameter C_S of the Smagorinsky models (8.277). Altogether, the use of the bubble VMS method is not recommended in John and Kindl (2010a). □

8.8.5 *Three-Scale Algebraic Variational Multiscale-Multigrid Methods (AVM^3 and AVM^4)*

Remark 8.252 (History) The AVM^3 method was introduced in Gravemeier et al. (2009, 2010) and developed further in Rasthofer and Gravemeier (2013) to AVM^4. □

Remark 8.253 (The Definition of the Small Resolved Scales) The definition of the small resolved scales uses an idea from algebraic multigrid (AMG) methods. The motivation for this approach comes from the goal to define the scale separation of the resolved scales without introducing another finite element space or another grid.

AMG methods are a proposal for transferring the ideas of geometric multigrid methods, see Remark 9.11, to problems where coarser geometric grids are not available, e.g., see Stüben (2001) for a description and a review. To this end, a multilevel structure is constructed solely based on a matrix, which represents the problem on the given grid. Then, coarser levels, discrete operators on these levels, and transfer operators (restriction and prolongation) are constructed. For the scale separation in AVM^3, only the construction of one coarser level and the corresponding transfer operators is needed.

There are several possibilities for constructing coarser levels in AMG methods. For AVM^3, it is proposed to use the simplest one, namely plain aggregation, see Vaněk et al. (1996). The degrees of freedom on the given grid correspond to the rows of the given matrix A. In Gravemeier et al. (2009, 2010) some root degree of freedom i is chosen and an aggregate is formed from the union of all degrees of freedom j for which the matrix entry a_{ij} does not vanish. Then, these degrees of freedom are removed from the list, a next root degree of freedom is chosen and this procedure is continued until all degrees of freedom belong to an aggregate. There are also other possibilities for choosing the aggregates, e.g., based on the strength of the coupling in the matrix A, i.e., based on $\left|a_{ij}\right|$. The aggregates represent the degrees of freedom on the coarse level. Denoting the fine and the coarse level in terms of the mesh width h of the geometric grid corresponding to the fine level, then the aggregates on the coarse level are usually denoted by $3h$.

Next, operators for the restriction R_h^{3h} and the prolongation of P_{3h}^h of vectors have to be defined. To this end, consider the matrix $\tilde{A}$ which differs from A only in the way that essential boundary conditions are replaced with natural boundary conditions. Let $\tilde{A}_0$ be a matrix whose columns span the kernel of $\tilde{A}$, i.e.,

$$\tilde{A}\tilde{A}_0 = 0. \tag{8.278}$$

The matrix on the coarse grid can be defined with the so-called Galerkin projection

$$\tilde{A}^{3h} = R_h^{3h}\tilde{A}P_{3h}^h.$$

Denoting the matrix which spans the kernel of $\tilde{A}^{3h}$ by $\tilde{A}_0^{3h}$, one obtains

$$0 = \tilde{A}^{3h}\tilde{A}_0^{3h} = R_h^{3h}\tilde{A}P_{3h}^h\tilde{A}_0^{3h}.$$

With (8.278), it follows that this equation is satisfied if

$$P_{3h}^h\tilde{A}_0^{3h} = \tilde{A}_0. \tag{8.279}$$

Based on (8.279), the operators P_{3h}^h and $\tilde{A}_0^{3h}$ can be determined simultaneously, see Gravemeier et al. (2009) for details. Finally, one sets

$$R_h^{3h} = \left(P_{3h}^h\right)^T.$$

Note that these operators are linear operators between finite-dimensional spaces and thus they can be represented with matrices.

Now, the operator for defining the large scales is given by

$$S_h^{3h} \ : \ V^h \to V^h, \quad \boldsymbol{u}^{3h} = P_{3h}^h R_h^{3h} \boldsymbol{u}^h,$$

i.e., in the first step $\boldsymbol{u}^h$ is restricted to the aggregates and in the second step, the representation of the aggregates in the finite element space is obtained. The small resolved scales are defined by

$$\boldsymbol{u}^h = \boldsymbol{u}^{3h} + \hat{\boldsymbol{u}}^h \quad \Longleftrightarrow \quad \hat{\boldsymbol{u}}^h = \boldsymbol{u}^h - \boldsymbol{u}^{3h}. \tag{8.280}$$

In AVM^3 from Gravemeier et al. (2010), the definition of the aggregates is based on the matrix that contains the complete discretization of the velocity-velocity part of the Navier–Stokes equations, including terms coming from stabilizations. □

*Remark 8.254 (*AVM^3*)* The derivation of this method can be explained by considering first a two-scale decomposition of the velocity and pressure

$$\boldsymbol{u} = \boldsymbol{u}^h + \boldsymbol{u}', \quad p = p^h + p', \tag{8.281}$$

where $(\boldsymbol{u}^h, p^h) \in V^h \times Q^h$ and V^h, Q^h are conforming finite element spaces. Then, the equation with the unresolved test functions is neglected. For the equation with the test functions from the finite element spaces, one obtains, using the decomposition (8.281),

$$\begin{aligned}
&\left(\partial_t \boldsymbol{u}^h, \boldsymbol{v}^h\right) + \left(2\nu \mathbb{D}\left(\boldsymbol{u}^h\right), \mathbb{D}\left(\boldsymbol{v}^h\right)\right) + \left(\left(\boldsymbol{u}^h \cdot \nabla\right) \boldsymbol{u}^h, \boldsymbol{v}^h\right) + \left(\nabla \cdot \boldsymbol{u}^h, q^h\right) \\
&\quad - \left(\nabla \cdot \boldsymbol{v}^h, p^h\right) \qquad\qquad (8.282)\\
&= \left(\boldsymbol{f}, \boldsymbol{v}^h\right) - \Big[\left(\partial_t \boldsymbol{u}', \boldsymbol{v}^h\right) + \left(2\nu \mathbb{D}\left(\boldsymbol{u}'\right), \mathbb{D}\left(\boldsymbol{v}^h\right)\right) + \left(\left(\boldsymbol{u}^h \cdot \nabla\right) \boldsymbol{u}', \boldsymbol{v}^h\right) \\
&\quad + \left(\left(\boldsymbol{u}' \cdot \nabla\right) \boldsymbol{u}^h, \boldsymbol{v}^h\right) + \left(\left(\boldsymbol{u}' \cdot \nabla\right) \boldsymbol{u}', \boldsymbol{v}^h\right) - \left(\nabla \cdot \boldsymbol{v}^h, p'\right) \Big] - \left(\nabla \cdot \boldsymbol{u}', q^h\right).
\end{aligned}$$

Now, the term in the brackets is considered and the test function is split into $\boldsymbol{v}^h = \boldsymbol{v}^{3h} + \hat{\boldsymbol{v}}^h$ in this term. Then, the assumption for a three-scale VMS method are applied, see Remark 8.214:

- The direct impact of the unresolved scales on the large scales is negligible, i.e., all terms in the brackets with test function $\boldsymbol{v}^{3h}$ are neglected.
- The direct impact of the unresolved scales onto the small resolved scales is modeled with a turbulence model, i.e., all terms in the brackets with test function $\hat{\boldsymbol{v}}^h$ are modeled. In Gravemeier et al. (2010), a Smagorinsky model of the form

$$\nabla \cdot \left(C_S h^2 \left\| \mathbb{D}\left(\hat{\boldsymbol{u}}^h\right) \right\|_{\mathrm{F}} \mathbb{D}\left(\hat{\boldsymbol{u}}^h\right)\right) = \nabla \cdot \left(\nu_{\mathrm{T}}\left(\hat{\boldsymbol{u}}^h\right) \mathbb{D}\left(\hat{\boldsymbol{u}}^h\right)\right) \tag{8.283}$$

was used.

Thus, the model of the term in the brackets in (8.282) reduces to (8.283). It contains the deformation tensor of the small resolved scales. □

Remark 8.255 (Realization of AVM^3*)* A realization of AVM^3 can be found so far only for the Q_1/Q_1 pair of finite element spaces. To account for the violation of the discrete inf-sup condition in this case, it was proposed in Gravemeier et al. (2010) to include a consistent stabilization that includes the PSPG stabilization, see Sect. 4.5.1, as model of the last term in (8.282)

$$\left(\nabla \cdot \boldsymbol{w}', q^h\right) \approx \sum_{K \in \mathcal{T}^h} \left(\partial_t \boldsymbol{w}^h - \nu \Delta \boldsymbol{w}^h + \left(\boldsymbol{w}^h \cdot \nabla\right) \boldsymbol{w}^h + \nabla r^h - \boldsymbol{f}, \delta_K^p \nabla q^h\right)_K,$$

where the convention for the notation given in Remark 8.217 was used. Then, the continuous-in-time AVM^3 method reads as follows: Find $\boldsymbol{w}^h : (0, T] \to V^h$, $r^h : (0, T] \to Q^h$ satisfying

$$\begin{aligned}
&\left(\partial_t \boldsymbol{w}^h, \boldsymbol{v}^h\right) + \left(2\nu \mathbb{D}\left(\boldsymbol{w}^h\right), \mathbb{D}\left(\boldsymbol{v}^h\right)\right) + \left(\left(\boldsymbol{w}^h \cdot \nabla\right) \boldsymbol{w}^h, \boldsymbol{v}^h\right) \\
&\quad + \left(\nabla \cdot \boldsymbol{w}^h, q^h\right) - \left(\nabla \cdot \boldsymbol{v}^h, r^h\right) + \left(\nu_T\left(\hat{\boldsymbol{w}}^h\right) \mathbb{D}\left(\hat{\boldsymbol{w}}^h\right), \mathbb{D}\left(\hat{\boldsymbol{v}}^h\right)\right) \\
&\quad + \sum_{K \in \mathcal{T}^h} \left(\partial_t \boldsymbol{w}^h - \nu \Delta \boldsymbol{w}^h + \left(\boldsymbol{w}^h \cdot \nabla\right) \boldsymbol{w}^h + \nabla r^h, \delta_K^p \nabla q^h\right)_K, \\
&= \left(\boldsymbol{f}, \boldsymbol{v}^h\right) + \sum_{K \in \mathcal{T}^h} \left(\boldsymbol{f}, \delta_K^p \nabla q^h\right)_K, \qquad (8.284)
\end{aligned}$$

where $\hat{\boldsymbol{w}}^h$ is computed with the help of the AMG approach sketched in Remark 8.253.

Using the short form (8.243), the method AVM^3 can be written as follows

$$\begin{aligned}
&A\left(\boldsymbol{w}^h; \left(\boldsymbol{w}^h, r^h\right), \left(\boldsymbol{v}^h, q^h\right)\right) + \text{PSPG-type stabilization} \qquad (8.285) \\
&\qquad + \left(\nu_T\left(\hat{\boldsymbol{w}}^h\right) \mathbb{D}\left(\hat{\boldsymbol{w}}^h\right), \mathbb{D}\left(\hat{\boldsymbol{v}}^h\right)\right) = F\left(\boldsymbol{v}^h\right).
\end{aligned}$$

□

Remark 8.256 (Numerical Experience) The algebraic VMS method AVM^3 was compared in Gravemeier et al. (2010) with the two-scale residual-based VMS method from Bazilevs et al. (2007) presented in Sect. 8.8.2. A turbulent channel flow problem, as described in Example D.12, and a turbulent lid driven cavity problem were considered. The simulations were performed for Q_1/Q_1 finite elements. It was observed that the results with AVM^3 were more accurate in several aspects and the simulations were somewhat more efficient.

Only small differences in accuracy and efficiency between both VMS methods were observed in Gravemeier et al. (2011) for the simulation of a turbulent flow around a cylinder, see Example D.13. Both VMS methods turned out to be clearly more efficient than the dynamic Smagorinsky model described in Remark 8.126.

□

Remark 8.257 (Multifractal Model of the Unresolved Scales—the Algebraic Multiscale-Multigrid-Multifractal Method AVM^4*)* In Rasthofer and Gravemeier (2013), it is proposed to use a so-called multifractal model of $\boldsymbol{u}'$ in (8.282) instead of modeling the terms in the brackets all together with an eddy viscosity model. Multifractal modeling of the unresolved scales is based on physical considerations, see Burton and Dahm (2005a,b) for a detailed derivation. As final result, the unresolved velocity scales can be represented in the form

$$\boldsymbol{u}' = C_{\mathrm{sgs}} \left(1 - \alpha^{-4/3}\right)^{-1/2} 2^{-2N/3} \left(2^{4N/3} - 1\right)^{1/2} \hat{\boldsymbol{u}}^h, \tag{8.286}$$

see Rasthofer and Gravemeier (2013). In (8.286), C_{sgs} is a constant, the parameter α comes from the definition of the large scales $\boldsymbol{u}^{\alpha h}$, i.e., $\alpha = 3$ in (8.280), and

$$N = \log_2 \left(\frac{h_K}{\lambda_\nu}\right) \tag{8.287}$$

is the number of cascades, which depends on the local mesh width and the viscous scale length λ_ν, see Remark 8.13. Model (8.286) is inserted in (8.282).

In Rasthofer and Gravemeier (2013), the value $C_{\mathrm{sgs}} = 0.25$ was used. The viscous scale length is about six times larger than the Kolmogorov scale (8.3), Rasthofer and Gravemeier (2013). The following approximations were proposed in Rasthofer and Gravemeier (2013), Rasthofer (2015, Sect. 4.2.5)

$$\frac{h_K}{\lambda_\nu} = C_\nu \left(\mathrm{Re}_K^h\right)^{3/4}$$

with $C_\nu = 1/12.3$ or $C_\nu = 1/11.2$ and

$$\mathrm{Re}_K^h = \frac{\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{\mathrm{F}} h_K^2}{\nu} \quad \text{or} \quad \mathrm{Re}_K^h = \frac{\left\|\boldsymbol{w}^h\right\|_2 h_K}{\nu}.$$

With these definitions, the value obtained on the right-hand side of (8.287) is usually not a natural number (as the notation "number of steps" would suggest). In practice, the generally non-natural numbers that are computed with the right-hand side of (8.287) are used for N, which can be seen, e.g., in Rasthofer and Gravemeier (2013, Fig. 11) or Rasthofer (2015, Fig. 4.7).

The multifractal modeling can be adapted to wall-bounded turbulent flows and it allows backscatter, see Rasthofer and Gravemeier (2013) and Rasthofer (2015) for details. To enhance numerical stability, it is proposed in Rasthofer and Gravemeier (2013) and Rasthofer (2015) to extend the multifractal model with residual-based stabilization terms, namely the SUPG term, the grad-div term, and the PSPG term. The arising method is called algebraic Multiscale-multigrid-multifractal method AVM^4 in Rasthofer (2015).

The method AVM^4 is compared in Rasthofer (2015) with the two-scale residual-based VMS method from Bazilevs et al. (2007), see Sect. 8.8.2, and the dynamic

Smagorinsky model presented in Remark 8.126. The simulations were performed with the Q_1/Q_1 pair of finite element spaces. It was shown that the adaption at the wall that is described in Rasthofer (2015) is of great importance for computing accurate solutions. For turbulent channel flows, see Example D.12, substantial better results were obtained with AVM^4 compared with the other methods. Also for the turbulent flow around a cylinder, Example D.13, AVM^4 gave the best results near the cylinder. The computing times of AVM^4 and the residual-based VMS method were similar. □

8.8.6 A Three-Scale Coarse Space Projection-Based VMS Method

Remark 8.258 (Basic Idea of the Method) The basic idea of the method consists in introducing an eddy viscosity term where the eddy viscosity does not act directly on all resolved scales but only on small resolved scales. This idea is the same as for the algebraic VMS methods described in Sect. 8.8.5. In the three-scale coarse space projection-based VMS method, the scale separation of the large scales is performed with the $L^2(\Omega)$ projection of the deformation tensor of the resolved scale velocity. □

8.8.6.1 Definition of the Method

Remark 8.259 (A Coarse Space Projection-Based VMS Method) Let $V^h \times Q^h$ be finite element spaces for the velocity and pressure that satisfy the discrete inf-sup stability condition (3.51), let L^H be a finite-dimensional space of symmetric $d \times d$ tensor-valued functions defined on Ω and let $\nu_{\mathrm{T}}\left(\left(\boldsymbol{w}^h, r^h\right), h\right)$ be a non-negative function. Then, the semi-discrete coarse space projection-based VMS method (continuous-in-time) is defined as follows: Find $\boldsymbol{w}^h : (0,T] \to V^h$, $r^h : (0,T] \to Q^h$, and $\mathbb{G}^H : (0,T] \to L^H$ satisfying

$$\begin{aligned}
\left(\partial_t \boldsymbol{w}^h, \boldsymbol{v}^h\right) + \left(2\nu\mathbb{D}\left(\boldsymbol{w}^h\right), \mathbb{D}\left(\boldsymbol{v}^h\right)\right) + n(\boldsymbol{w}^h, \boldsymbol{w}^h, \boldsymbol{v}^h)& \\
- \left(\nabla\cdot\boldsymbol{v}^h, r^h\right) + \left(\nu_{\mathrm{T}}\left(\mathbb{D}\left(\boldsymbol{w}^h\right) - \mathbb{G}^H\right), \mathbb{D}\left(\boldsymbol{v}^h\right)\right) &= \langle \boldsymbol{f}, \boldsymbol{v}^h\rangle_{V',V}, \\
\left(\nabla\cdot\boldsymbol{w}^h, q^h\right) &= 0, \\
\left(\mathbb{D}\left(\boldsymbol{w}^h\right) - \mathbb{G}^H, \mathbb{L}^H\right) &= 0,
\end{aligned} \tag{8.288}$$

for all $\left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h$ and $\mathbb{L}^H \in L^H$. The scales are defined by projection in the last equation of (8.288), there are large scales and small resolved scales, and the turbulence model is applied directly only to the small resolved scales, see Remark 8.261 for a detailed discussion of the last two issues.

Using the short form (8.243), the first two equations of (8.288) can be written in the form

$$A\left(\boldsymbol{w}^h;\left(\boldsymbol{w}^h,r^h\right),\left(\boldsymbol{v}^h,q^h\right)\right)+\left(\nu_{\mathrm{T}}\left(\mathbb{D}\left(\boldsymbol{w}^h\right)-\mathbb{G}^H\right),\mathbb{D}\left(\boldsymbol{v}^h\right)\right)=F\left(\boldsymbol{v}^h\right). \tag{8.289}$$

Comparing this representation with (8.285) shows that, apart of the PSPG-type stabilization, the coarse space projection-based VMS method and AVM[3] have principally the same form. Only, in (8.289), the small resolved scales of the deformation tensor appear whereas (8.285) contains the deformation tensor of the small resolved scales. For constant turbulent viscosity ν_{T}, the second term on the left-hand side of (8.289) can be rewritten in terms of the deformation tensor of the small resolved scales, compare Remark 8.264.

The method (8.288) was proposed in John and Kaya (2005) based on ideas from Layton (2002). For applying this method, one has to choose two parameters: the additional viscosity $\nu_{\mathrm{T}}\left(\left(\boldsymbol{w}^h,r^h\right),h\right)$ and the space L^H. □

Remark 8.260 (Choice of the Additional Viscosity) Concerning the turbulent viscosity $\nu_{\mathrm{T}}\left(\left(\boldsymbol{w}^h,r^h\right),h\right)$, numerical studies with the method (8.288) presented in John and Kaya (2005), John and Roland (2007), John and Kindl (2010a,b) and Röhe and Lube (2010) used a Smagorinsky models of the form

$$\nu_{\mathrm{T}}=C_S\delta^2\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{\mathrm{F}}, \tag{8.290}$$

$$\nu_{\mathrm{T}}=C_S\delta^2\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)-\mathbb{G}^H\right\|_{\mathrm{F}}, \tag{8.291}$$

$$\nu_{\mathrm{T}}|_K=C_S\frac{\delta^2}{|K|^{1/2}}\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)-\mathbb{G}^H\right\|_{L^2(K)}. \tag{8.292}$$

Altogether, the typical feature of a three-scale VMS method, namely that the turbulence model is applied directly only to the small resolved scales, can be observed very well in the last term on the left-hand side of the first equation of (8.288). □

Remark 8.261 (Choice of the Large Scale Projection Space) The other parameter in (8.288) is the space of symmetric tensors L^H. The last equation in (8.288) states that the tensor $\mathbb{G}^H$ is just the $L^2(\Omega)$ projection of $\mathbb{D}\left(\boldsymbol{w}^h\right)$ into L^H: $P_{L^H}\ :\ L=\mathbb{D}(V)\to L^H$, $\mathbb{D}(\boldsymbol{v})\to P_{L^H}\mathbb{D}(\boldsymbol{v})=\mathbb{G}^H$ with

$$\left(P_{L^H}\mathbb{D}(\boldsymbol{v})-\mathbb{D}(\boldsymbol{v}),\mathbb{L}^H\right)=0\quad\forall\,\mathbb{L}^H\in L^H. \tag{8.293}$$

With this notation, one can reformulate the short form (8.289) as follows: Find $\boldsymbol{w}^h\ :\ (0,T]\to V^h,\ r^h\ :\ (0,T]\to Q^h$ satisfying

$$A\left(\boldsymbol{w}^h;\left(\boldsymbol{w}^h,r^h\right),\left(\boldsymbol{v}^h,q^h\right)\right)+\left(\nu_{\mathrm{T}}\left(I-P_{L^H}\right)\mathbb{D}\left(\boldsymbol{w}^h\right),\mathbb{D}\left(\boldsymbol{v}^h\right)\right)=\langle\boldsymbol{f},\boldsymbol{v}^h\rangle_{V',V} \tag{8.294}$$

for all $\left(\boldsymbol{v}^h,q^h\right)\in V^h\times Q^h$.

The space L^H plays the role of a large scale space, i.e., $(I - P_{L^H})\,\mathbb{D}\left(\boldsymbol{w}^h\right)$ represents small resolved scales of $\mathbb{D}\left(\boldsymbol{w}^h\right)$. To avoid a negative additional viscosity, it is required that $L^H \subset \{\mathbb{D}\left(\boldsymbol{v}^h\right) \,:\, \boldsymbol{v}^h \in V^h\}$.

Considering the limit cases for L^H gives the following models.

- In the case that both spaces are identical, the second term on the left-hand side of (8.294) vanishes and the Galerkin finite element discretization of the Navier–Stokes equations is recovered.
- If $L^H = \{\mathbb{O}\}$, one obtains an artificial viscosity stabilization of the Navier–Stokes equations with a possible nonlinear artificial viscosity. If $\nu_{\mathrm{T}}\left(\left(\boldsymbol{w}^h, r^h\right), h\right)$ is the Smagorinsky eddy viscosity model (8.67), the Smagorinsky LES model is recovered.

Since L^H represents large scales, it has to be in some sense a coarse finite element space. There are essentially two possibilities:

- If V^h is a higher order finite element space, L^H can be defined as low order finite element space on the same grid as V^h. This approach was developed in John and Kaya (2005) and it will be discussed in this section.
- The second possibility, in particular if V^h is a low order discretization, consists in defining L^H on a coarser grid, see John et al. (2006b) for a study of this approach in the case of convection-dominated convection-diffusion equations.

Since $\mathbb{D}\left(\boldsymbol{w}^h\right)$ is a discontinuous piecewise polynomial tensor, choosing its $L^2(\Omega)$ projection in the same way seems to be natural. Thus, L^H should consist of discontinuous piecewise polynomial tensors. It will be explained in Remark 8.275 that this choice is mandatory for the sake of an efficient implementation. □

8.8.6.2 Imbedding the Method into the Basic Approach From Sect. 8.8.1

Remark 8.262 (Coarse Spaces for Velocity and Pressure and Corresponding Projections) The method (8.288) can be transformed, in the case ν_{T} being a positive constant, to the abstract form (8.244), (8.245) of a VMS method. To this end, the three-scale partitioning given in Remark 8.214 has to be described by appropriately chosen function spaces and projections.

Clearly, the continuous pair of spaces $V \times Q$ contains all scales. The finite element spaces $V^h \times Q^h$ contain the large and the small resolved scales.

Let $V^H \subset H^1(\Omega)$ be a discrete space such that $L^H = \mathbb{D}\left(V^H\right)$. The space V^H should be coarser than V^h. But in the definition of V^H, no essential boundary conditions, like no-slip conditions, are incorporated. Thus, in general $V^H \not\subset V^h$. The pair of spaces for the large scales is given by $V^H \times Q^H$, where Q^H is chosen such that a discrete inf-sup condition of type (3.51) is fulfilled for $V^H \times Q^H$. Then, the large scales $P_H\boldsymbol{u}$ of the velocity are defined by an elliptic projection into V^H (with natural boundary conditions) and the large scales $P_H p$ of the pressure by the

$L^2(\Omega)$ projection into Q^H: $P_H \ : \ V \times Q \to V^H \times Q^H$

$$\begin{aligned} \left(\mathbb{D}\left(\boldsymbol{u} - P_H\boldsymbol{u}\right), \mathbb{D}\left(\overline{\boldsymbol{v}}^H\right)\right) &= 0 \quad \forall\ \overline{\boldsymbol{v}}^H \in V^H, \\ \left(\boldsymbol{u} - P_H\boldsymbol{u}, 1\right) &= 0, \\ \left(p - P_H p, q^H\right) &= 0 \quad \forall\ q^H \in Q^H. \end{aligned} \tag{8.295}$$

□

Lemma 8.263 (Commutation of the Definition of the Large Scales and Differentiation) *Let $\boldsymbol{v} \in V$, $L^H = \mathbb{D}\left(V^H\right)$ and denote by $P_{L^H}\mathbb{D}\left(\boldsymbol{v}\right)$ the $L^2(\Omega)$ projection of $\mathbb{D}\left(\boldsymbol{v}\right)$ into L^H defined in the last equation of* (8.288)*. Then*

$$P_{L^H}\mathbb{D}\left(\boldsymbol{v}\right) = \mathbb{D}\left(P_H\boldsymbol{v}\right) \quad \forall\ \boldsymbol{v} \in V, \tag{8.296}$$

where for simplicity of notation, the symbol P_H is used just for the velocity part of the map defined in (8.295)*.*

Proof From $L^H = \mathbb{D}\left(V^H\right)$ and $P_{L^H}\mathbb{D}\left(\boldsymbol{v}\right) \in L^H$, it follows that there is a $\boldsymbol{w}^H \in V^H$ such that $P_{L^H}\mathbb{D}\left(\boldsymbol{v}\right) = \mathbb{D}\left(\boldsymbol{w}^H\right)$. Using the last equation of (8.288) gives

$$\left(\mathbb{D}\left(\boldsymbol{v} - \boldsymbol{w}^H\right), \mathbb{L}^H\right) = 0 \quad \forall\ \mathbb{L}^H \in L^H. \tag{8.297}$$

On the other hand, since $L^H = \mathbb{D}\left(V^H\right)$, (8.295) is equivalent to

$$\left(\mathbb{D}\left(\boldsymbol{v} - P_H\boldsymbol{v}\right), \mathbb{L}^H\right) = 0 \quad \forall\ \mathbb{L}^H \in L^H. \tag{8.298}$$

The statement of the lemma follows now directly from (8.297) and (8.298) since the elliptic projection is unique. ■

Remark 8.264 (Transform to the Form (8.244)*,* (8.245) *for a Constant Additional Viscosity)* Let ν_{T} be a positive constant. Since $P_{L^H}\mathbb{D}\left(\boldsymbol{v}^h\right) \in L^H$, it follows from the last equation of (8.288) that

$$\left(\left(I - P_{L^H}\right)\mathbb{D}\left(\boldsymbol{w}^h\right), P_{L^H}\mathbb{D}\left(\boldsymbol{v}^h\right)\right) = 0$$

and consequently, one gets for a constant viscosity ν_{T}

$$\left(\nu_{\mathrm{T}}\left(I - P_{L^H}\right)\mathbb{D}\left(\boldsymbol{w}^h\right), \mathbb{D}\left(\boldsymbol{v}^h\right)\right) = \left(\nu_{\mathrm{T}}\left(I - P_{L^H}\right)\mathbb{D}\left(\boldsymbol{w}^h\right), \left(I - P_{L^H}\right)\mathbb{D}\left(\boldsymbol{v}^h\right)\right).$$

Thus, (8.289) can be reformulated as follows: Find $\boldsymbol{w}^h \ : \ (0, T] \to V^h$, $r^h \ : \ (0, T] \to Q^h$ satisfying

$$\begin{aligned} &A\left(\boldsymbol{w}^h; \left(\boldsymbol{w}^h, r^h\right), \left(\boldsymbol{v}^h, q^h\right)\right) \\ &+ \left(\nu_{\mathrm{T}}\left(I - P_{L^H}\right)\mathbb{D}\left(\boldsymbol{w}^h\right), \left(I - P_{L^H}\right)\mathbb{D}\left(\boldsymbol{v}^h\right)\right) = \langle \boldsymbol{f}, \boldsymbol{v}^h \rangle_{V',V} \end{aligned} \tag{8.299}$$

for all $(\boldsymbol{v}^h, q^h) \in V^h \times Q^h$. Decomposing $V^h = V^H + \widehat{V}^h$ and $Q^h = Q^H + \widehat{Q}^h$ with $\widehat{V}^h = (I - P_H)\, V^h$, one obtains with (8.296)

$$(I - P_{L^H})\, \mathbb{D}\left(\boldsymbol{v}^h\right) = \mathbb{D}\left(\boldsymbol{v}^h - P_H \boldsymbol{v}^h\right) = \mathbb{D}\left((I - P_H)\, \boldsymbol{v}^h\right) = \mathbb{D}\left(\hat{\boldsymbol{v}}^h\right).$$

The decompositions $\boldsymbol{w}^h = \overline{\boldsymbol{u}}^H + \hat{\boldsymbol{w}}^h$, $r^h = \overline{p}^H + \hat{p}^H$, $\boldsymbol{v}^h = \overline{\boldsymbol{v}}^H + \hat{\boldsymbol{v}}^h$, and $q^h = \overline{q}^H + \hat{q}^H$ are inserted in (8.299). Using the linearity of $A(\cdot;\cdot,\cdot)$ with respect to the second and third component and writing the arising equation formally as a coupled system gives

$$\begin{aligned} & A\left(\overline{\boldsymbol{u}}^H + \hat{\boldsymbol{w}}^h; \left(\overline{\boldsymbol{u}}^H, \overline{p}^H\right), \left(\overline{\boldsymbol{v}}^H, \overline{q}^H\right)\right) \\ & + A\left(\overline{\boldsymbol{u}}^H + \hat{\boldsymbol{w}}^h; \left(\hat{\boldsymbol{w}}^h, \hat{r}^h\right), \left(\overline{\boldsymbol{v}}^H, \overline{q}^H\right)\right) = F\left(\overline{\boldsymbol{v}}^H\right) \end{aligned} \tag{8.300}$$

for all test functions $\left(\overline{\boldsymbol{v}}^H, \overline{q}^H\right) \in V^H \times Q^H$ and

$$\begin{aligned} & A\left(\overline{\boldsymbol{u}}^H + \hat{\boldsymbol{w}}^h; \left(\overline{\boldsymbol{u}}^H, \overline{p}^H\right), \left(\hat{\boldsymbol{v}}^h, \hat{q}^h\right)\right) \\ & + A\left(\overline{\boldsymbol{u}}^H + \hat{\boldsymbol{w}}^h; \left(\hat{\boldsymbol{w}}^h, \hat{r}^h\right), \left(\hat{\boldsymbol{v}}^h, \hat{q}^h\right)\right) + \left(\nu_{\mathrm{T}} \mathbb{D}\left(\hat{\boldsymbol{w}}^h\right), \mathbb{D}\left(\hat{\boldsymbol{v}}^h\right)\right) = F\left(\hat{\boldsymbol{v}}^h\right) \end{aligned} \tag{8.301}$$

for all test functions from $\widehat{V}^h \times \widehat{Q}^h$. The coupled system (8.300), (8.301) possesses exactly the form (8.244), (8.245). The unresolved scales are modeled only in the small scale equation (8.301) with the model

$$T\left(\boldsymbol{w}^h; \left(\overline{\boldsymbol{u}}^H, \overline{p}^H\right), \left(\hat{\boldsymbol{u}}^h, \hat{p}^h\right), \left(\hat{\boldsymbol{v}}^h, \hat{q}^h\right)\right) = \left(\nu_{\mathrm{T}} \mathbb{D}\left(\hat{\boldsymbol{w}}^h\right), \mathbb{D}\left(\hat{\boldsymbol{v}}^h\right)\right)$$

and this model influences the large scales solely indirectly by the coupling of (8.300) and (8.301). □

8.8.6.3 Finite Element Error Analysis

Remark 8.265 (Finite Element Error Analysis in the Literature) A finite element error analysis for different versions of the three-scale projection-based VMS method can be found in John and Kaya (2008), John et al. (2008) and Röhe and Lube (2010). The analysis was always performed for the continuous-in-time case.

- In John and Kaya (2008), the case of a constant viscosity ν_{T} was considered.
- The case of ν_{T} being of Smagorinsky type, using the small resolved scales in the definition of the Smagorinsky term, was studied in John et al. (2008). In this paper, the additional viscous term was defined differently than in (8.288). There, it is the deformation tensor of the small resolved scales and not the small resolved scales of the deformation tensor, i.e., differentiation and projection were

interchanged. With this interchange, one gets an additional viscous term in the momentum equation that has the same form as the Smagorinsky term (8.67), but only with the small resolved scales instead with all resolved scales. Then the analysis follows the lines of Sect. 8.3.3, using the same function spaces.

- The numerical analysis in Röhe and Lube (2010) considered the case of ν_T being a piecewise constant. In the discrete equations, also a grad-div stabilization term, see Sect. 4.6.1, was introduced. Besides an estimate for the velocity error, also an estimate for the pressure error was given.

□

Remark 8.266 (Turbulent Viscosity in the Analysis and Goal of the Analysis) Here, the analysis for ν_T being a piecewise constant function, for the discrete equations (8.288), i.e., without grad-div term, and for the continuous-in-time case will be presented. The case of ν_T being piecewise constant is much closer to the use of the projection-based VMS method in practice than a global constant ν_T. Neglecting the grad-div term serves for concentrating on the effect of the additional viscous term on the error bound. In Röhe and Lube (2010), a number of terms were estimated with the help of the grad-div stabilization term. It is also known that the grad-div term alone might give error estimates independent of the viscosity, see de Frutos et al. (2016b) for the time-dependent Oseen equations, and turbulent flow simulations might be stabilized strongly by using only this term, see John and Kindl (2010a).

The analysis follows John and Kaya (2008), Röhe and Lube (2010) and the goal consists in proving an error estimate where some constants depend on inverse powers of a modified viscosity and not, as for the Galerkin finite element method, on inverse powers of the viscosity, see estimate (7.44). The modified viscosity is not smaller than ν. □

Remark 8.267 (The Continuous Equation) The continuous Navier–Stokes equations will be considered in the deformation tensor form Find $\boldsymbol{u} : (0,T] \to V$ and $p : (0,T] \to Q$ such that

$$(\partial_t \boldsymbol{u}, \boldsymbol{v}) + (2\nu\mathbb{D}(\boldsymbol{u}), \mathbb{D}(\boldsymbol{v})) + n_{\text{skew}}(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}) - (\nabla\cdot\boldsymbol{v}, p) + (\nabla\cdot\boldsymbol{u}, q) = \langle \boldsymbol{f}, \boldsymbol{v}\rangle_{V',V} \tag{8.302}$$

for all $(\boldsymbol{v}, q) \in V \times Q$ and $\boldsymbol{u}(0,\boldsymbol{x}) = \boldsymbol{u}_0(\boldsymbol{x}) \in H_{\text{div}}(\Omega)$. Note that this form is equivalent to (7.40).

Similarly to the proof of Lemma 7.21, using in addition Korn's inequality (3.43), the following stability estimate can be shown

$$\|\boldsymbol{u}(t)\|_{L^2(\Omega)}^2 + 2\nu\|\mathbb{D}(\boldsymbol{u})\|_{L^2(0,t;L^2(\Omega))}^2 \le \|\boldsymbol{u}(0)\|_{L^2(\Omega)}^2 + \frac{C}{\nu}\|\boldsymbol{f}\|_{L^2(0,t;H^{-1}(\Omega))}^2. \tag{8.303}$$

It follows that $\boldsymbol{u} \in L^\infty(0,T;L^2(\Omega)) \cap L^2(0,T;V_{\text{div}})$. □

Remark 8.268 (Definition of a Modified Viscosity) Let $\mathcal{T}^h$ be a triangulation and $\nu_T(\boldsymbol{w}^h)$ be piecewise constant with respect to $\mathcal{T}^h$. The turbulent viscosity might depend on the solution of (8.288). The restriction to a mesh cell $K \in \mathcal{T}^h$ will be denoted by $\nu_{\mathrm{T}}^{K}\left(\boldsymbol{w}^{h}\right)$. Define for all $\boldsymbol{v} \in V$

$$P_{L^H}^{K}\mathbb{D}(\boldsymbol{v}) = \begin{cases} P_{L^H}\mathbb{D}(\boldsymbol{v}) & \text{in } K, \\ \{\mathbb{O}\} & \text{else.} \end{cases} \tag{8.304}$$

Since L^H consists of discontinuous functions, it follows that $P_{L^H}^{K}\mathbb{D}(\boldsymbol{v}) \in L^H$. By definition, the support of $P_{L^H}^{K}\mathbb{D}(\boldsymbol{v})$ is just the mesh cell K. Using (8.304), $P_{L^H}^{K}\mathbb{D}(\boldsymbol{v}) \in L^H$, the last equation of (8.288), and again (8.304) yields for all $\boldsymbol{v} \in V$

$$\begin{aligned}
&\|(I - P_{L^H})\,\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2 \\
&= \|\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2 - 2\left(\mathbb{D}(\boldsymbol{v}), P_{L^H}\mathbb{D}(\boldsymbol{v})\right)_K + \|P_{L^H}\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2 \\
&= \|\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2 - 2\left(\mathbb{D}(\boldsymbol{v}), P_{L^H}^{K}\mathbb{D}(\boldsymbol{v})\right) + \|P_{L^H}\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2 \\
&= \|\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2 - 2\left(P_{L^H}\mathbb{D}(\boldsymbol{v}), P_{L^H}^{K}\mathbb{D}(\boldsymbol{v})\right) + \|P_{L^H}\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2 \\
&= \|\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2 - 2\left(P_{L^H}\mathbb{D}(\boldsymbol{v}), P_{L^H}\mathbb{D}(\boldsymbol{v})\right)_K + \|P_{L^H}\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2 \\
&= \|\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2 - \|P_{L^H}\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2 .
\end{aligned} \tag{8.305}$$

Using this Pythagorean identity, one gets for $\|\mathbb{D}(\boldsymbol{v})\|_{L^2(K)} > 0$

$$\begin{aligned}
&\sum_{K\in\mathcal{T}^h} \nu_{\mathrm{T}}^{K}\left(\boldsymbol{w}^{h}\right) \|(I - P_{L^H})\,\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2 \\
&\quad= \sum_{K\in\mathcal{T}^h} \nu_{\mathrm{T}}^{K}\left(\boldsymbol{w}^{h}\right) \left(\|\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2 - \|P_{L^H}\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2\right) \\
&\quad= \sum_{K\in\mathcal{T}^h} \nu_{\mathrm{T}}^{K}\left(\boldsymbol{w}^{h}\right) \left(1 - \frac{\|P_{L^H}\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2}{\|\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2}\right) \|\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2 \\
&\quad= \sum_{K\in\mathcal{T}^h} \nu_{\mathrm{VMS}}^{K}\left(\boldsymbol{w}^{h}, \boldsymbol{v}\right) \|\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2 .
\end{aligned} \tag{8.306}$$

One obtains with (8.304), the last equation of (8.288), again (8.304), and the Cauchy–Schwarz inequality (A.10)

$$\begin{aligned}
&\|P_{L^H}\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}^2 \\
&= \left(P_{L^H}\mathbb{D}(\boldsymbol{v}), P_{L^H}\mathbb{D}(\boldsymbol{v})\right)_K = \left(P_{L^H}\mathbb{D}(\boldsymbol{v}), P_{L^H}^{K}\mathbb{D}(\boldsymbol{v})\right) \\
&= \left(\mathbb{D}(\boldsymbol{v}), P_{L^H}^{K}\mathbb{D}(\boldsymbol{v})\right) = \left(\mathbb{D}(\boldsymbol{v}), P_{L^H}\mathbb{D}(\boldsymbol{v})\right)_K \\
&\le \|\mathbb{D}(\boldsymbol{v})\|_{L^2(K)} \|P_{L^H}\mathbb{D}(\boldsymbol{v})\|_{L^2(K)} .
\end{aligned}$$

Hence, it is $0 \le \|P_{L^H}\mathbb{D}(\boldsymbol{v})\|_{L^2(K)} \le \|\mathbb{D}(\boldsymbol{v})\|_{L^2(K)}$ and one concludes from the definition (8.306) of $\nu_{\mathrm{VMS}}^K\left(\boldsymbol{w}^h, \boldsymbol{v}\right)$ that

$$0 \le \nu_{\mathrm{VMS}}^K\left(\boldsymbol{w}^h, \boldsymbol{v}\right) \le \nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right). \tag{8.307}$$

Usually, $\boldsymbol{w}^h$ depends on time such that $\nu_{\mathrm{VMS}}^K\left(\boldsymbol{w}^h, \boldsymbol{v}\right)$ will be also time-dependent. In the case $\|\mathbb{D}(\boldsymbol{v})\|_{L^2(K)} = 0$, it is set $\nu_{\mathrm{VMS}}^K\left(\boldsymbol{w}^h, \boldsymbol{v}\right) = 0$. □

Remark 8.269 (Formulation of the Discrete Problem) Starting point for the formulation of the discrete problem is (8.288). Since $P_{L^H}^K\mathbb{D}(\boldsymbol{v}) \in L^H$ and the support of $P_{L^H}^K\mathbb{D}(\boldsymbol{v})$ is just K, one gets

$$\begin{aligned} 0 &= \left((I - P_{L^H})\,\mathbb{D}\left(\boldsymbol{w}^h\right), P_{L^H}^K\mathbb{D}(\boldsymbol{v})\right) = \left((I - P_{L^H})\,\mathbb{D}\left(\boldsymbol{w}^h\right), P_{L^H}^K\mathbb{D}(\boldsymbol{v})\right)_K \\ &= \left((I - P_{L^H})\,\mathbb{D}\left(\boldsymbol{w}^h\right), P_{L^H}\mathbb{D}(\boldsymbol{v})\right)_K. \end{aligned} \tag{8.308}$$

Using that $\nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right)$ is piecewise constant and (8.308) gives

$$\begin{aligned} &\left(\nu_{\mathrm{T}}\left(\mathbb{D}\left(\boldsymbol{w}^h\right) - G^H\right), \mathbb{D}\left(\boldsymbol{v}^h\right)\right) \\ &= \sum_{K\in\mathcal{T}^h} \nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right)\left((I - P_{L^H})\,\mathbb{D}\left(\boldsymbol{w}^h\right), \mathbb{D}\left(\boldsymbol{v}^h\right)\right)_K \\ &= \sum_{K\in\mathcal{T}^h} \nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right)\left((I - P_{L^H})\,\mathbb{D}\left(\boldsymbol{w}^h\right), (I - P_{L^H})\,\mathbb{D}\left(\boldsymbol{v}^h\right)\right)_K. \end{aligned}$$

In this way, the projection can be inserted in the momentum equation of (8.288). The discrete equation reads as follows: Find $\boldsymbol{w}^h : (0, T] \to V^h$, $r^h : (0, T] \to Q^h$ satisfying

$$\begin{aligned} &\left(\partial_t\boldsymbol{w}^h, \boldsymbol{v}^h\right) + \left(2\nu\mathbb{D}\left(\boldsymbol{w}^h\right), \mathbb{D}\left(\boldsymbol{v}^h\right)\right) + n_{\mathrm{skew}}\left(\boldsymbol{w}^h, \boldsymbol{w}^h, \boldsymbol{v}^h\right) - \left(\nabla\cdot\boldsymbol{v}^h, r^h\right) \\ &\qquad + \sum_{K\in\mathcal{T}^h} \nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right)\left((I - P_{L^H})\,\mathbb{D}\left(\boldsymbol{w}^h\right), (I - P_{L^H})\,\mathbb{D}\left(\boldsymbol{v}^h\right)\right)_K = \langle \boldsymbol{f}, \boldsymbol{v}^h\rangle_{V',V}, \\ &\qquad \left(\nabla\cdot\boldsymbol{w}^h, q^h\right) = 0, \end{aligned} \tag{8.309}$$

for all $\left(\boldsymbol{v}^h, q^h\right) \in V^h \times Q^h$, where the projection P_{L^H} was defined in (8.293).

In the finite element error analysis for the velocity, problem (8.309) is considered in V_{div}^h: Find $\boldsymbol{w}^h : (0, T] \to V_{\mathrm{div}}^h$ such that

$$\begin{aligned} &\left(\partial_t\boldsymbol{w}^h, \boldsymbol{v}^h\right) + \left(2\nu\mathbb{D}\left(\boldsymbol{w}^h\right), \mathbb{D}\left(\boldsymbol{v}^h\right)\right) + n_{\mathrm{skew}}\left(\boldsymbol{w}^h, \boldsymbol{w}^h, \boldsymbol{v}^h\right) \\ &+ \sum_{K\in\mathcal{T}^h} \nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right)\left((I - P_{L^H})\,\mathbb{D}\left(\boldsymbol{w}^h\right), (I - P_{L^H})\,\mathbb{D}\left(\boldsymbol{v}^h\right)\right)_K = \langle \boldsymbol{f}, \boldsymbol{v}^h\rangle_{V',V} \end{aligned} \tag{8.310}$$

for all $\boldsymbol{v}^h \in V_{\mathrm{div}}^h$. □

Remark 8.270 (Assumptions on the Data and the Solution of the Navier–Stokes Equations) For the finite element error analysis, some assumptions on the regularity of the solution and the data of the Navier–Stokes equations are needed. It will be assumed that

$$\boldsymbol{f} \in L^2\left(0,T;H^{-1}(\Omega)\right), \quad \boldsymbol{u}_0 \in H_{\text{div}}(\Omega), \tag{8.311}$$

and that (8.302) possesses a solution $(\boldsymbol{u}, p)$ with

$$\nabla\boldsymbol{u} \in L^4\left(0,T;L^2(\Omega)\right), \quad \partial_t\boldsymbol{u} \in L^2\left(0,T;H^{-1}(\Omega)\right), \quad p \in L^2\left(0,T;L^2(\Omega)\right). \tag{8.312}$$

□

Lemma 8.271 (Stability of $\boldsymbol{w}^h$) *Let Ω be a bounded domain with polyhedral and Lipschitz continuous boundary and let the regularity assumptions (8.311) hold. Assume that the three-scale projection-based VMS method (8.310) has a unique solution $\boldsymbol{w}^h$. Let for $\boldsymbol{v} \in V$*

$$\nu_{\text{mod}}^K\left(\boldsymbol{w}^h, \boldsymbol{v}\right) = 2\nu + \nu_{\text{VMS}}^K\left(\boldsymbol{w}^h, \boldsymbol{v}\right) \tag{8.313}$$

and

$$\nu_{\text{mod}}^{\min}\left(\boldsymbol{w}^h, \boldsymbol{v}\right) = \min_{K\in\mathcal{T}^h} \nu_{\text{mod}}^K\left(\boldsymbol{w}^h, \boldsymbol{v}\right). \tag{8.314}$$

Then, the solution $\boldsymbol{w}^h$ of (8.310) satisfies the stability estimate

$$\begin{aligned}\left\|\boldsymbol{w}^h(t)\right\|_{L^2(\Omega)}^2 &+ \int_0^t \sum_{K\in\mathcal{T}^h} \nu_{\text{mod}}^K\left(\boldsymbol{w}^h, \boldsymbol{w}^h\right) \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)(\tau)\right\|_{L^2(K)}^2 \, d\tau \\ &\le \left\|\boldsymbol{w}^h(0)\right\|_{L^2(\Omega)}^2 + C\int_0^t \frac{\|\boldsymbol{f}\|_{H^{-1}(\Omega)}^2}{\nu_{\text{mod}}^{\min}\left(\boldsymbol{w}^h, \boldsymbol{w}^h\right)} \, d\tau \end{aligned} \tag{8.315}$$

for all $t \in (0,T]$. It follows that $\boldsymbol{w}^h \in L^\infty\left(0,T;L^2(\Omega)\right)$ and $\mathbb{D}\left(\boldsymbol{w}^h\right) \in L^2\left(0,T;L^2(\Omega)\right)$.

Proof The stability estimate is derived as usual by using the solution as test function. One gets with (7.13), (8.306) for $\boldsymbol{v} = \boldsymbol{w}^h$, integration on $(0,t)$, the estimate for the dual pairing, the Poincaré inequality (A.12), Korn's inequality (3.43), and Young's inequality (A.5)

$$\begin{aligned}&\frac{1}{2}\left\|\boldsymbol{w}^h(t)\right\|_{L^2(\Omega)}^2 + \int_0^t \sum_{K\in\mathcal{T}^h} \nu_{\text{mod}}^K\left(\boldsymbol{w}^h, \boldsymbol{w}^h\right) \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)(\tau)\right\|_{L^2(K)}^2 \, d\tau \\ &= \frac{1}{2}\left\|\boldsymbol{w}^h(0)\right\|_{L^2(\Omega)}^2 + \int_0^t \left(\boldsymbol{f}, \boldsymbol{w}^h\right) \, d\tau \\ &\le \frac{1}{2}\left\|\boldsymbol{w}^h(0)\right\|_{L^2(\Omega)}^2 + C\int_0^t \|\boldsymbol{f}\|_{H^{-1}(\Omega)} \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^2(\Omega)} \, d\tau\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2}\left\|\boldsymbol{w}^h(0)\right\|_{L^2(\Omega)}^2 + C\int_0^t \|\boldsymbol{f}\|_{H^{-1}(\Omega)} \left(\sum_{K\in\mathcal{T}^h} \frac{\nu_{\mathrm{mod}}^K\left(\boldsymbol{w}^h,\boldsymbol{w}^h\right)}{\nu_{\mathrm{mod}}^K\left(\boldsymbol{w}^h,\boldsymbol{w}^h\right)} \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^2(K)}^2\right)^{1/2} d\tau \\
&\le \frac{1}{2}\left\|\boldsymbol{w}^h(0)\right\|_{L^2(\Omega)}^2 \\
&\quad + C\int_0^t \frac{\|\boldsymbol{f}\|_{H^{-1}(\Omega)}}{\left(\nu_{\mathrm{mod}}^{\min}\left(\boldsymbol{w}^h,\boldsymbol{w}^h\right)\right)^{1/2}} \left(\sum_{K\in\mathcal{T}^h} \nu_{\mathrm{mod}}^K\left(\boldsymbol{w}^h,\boldsymbol{w}^h\right) \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^2(K)}^2\right)^{1/2} d\tau \\
&\le \frac{1}{2}\left\|\boldsymbol{w}^h(0)\right\|_{L^2(\Omega)}^2 + C\int_0^t \frac{\|\boldsymbol{f}\|_{H^{-1}(\Omega)}^2}{\nu_{\mathrm{mod}}^{\min}\left(\boldsymbol{w}^h,\boldsymbol{w}^h\right)}\, d\tau \\
&\quad + \frac{1}{2}\int_0^t \sum_{K\in\mathcal{T}^h} \nu_{\mathrm{mod}}^K\left(\boldsymbol{w}^h,\boldsymbol{w}^h\right) \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^2(K)}^2\, d\tau.
\end{aligned} \tag{8.316}
$$

Absorbing the last term on the right-hand side in the left-hand side gives estimate (8.315).

The regularity $\boldsymbol{w}^h \in L^\infty\left(0,T;L^2(\Omega)\right)$ follows immediately from (8.315). Using $2\nu + \nu_{\mathrm{VMS}}^K\left(\boldsymbol{w}^h,\boldsymbol{w}^h\right) \ge 2\nu > 0$ gives also $\mathbb{D}\left(\boldsymbol{w}^h\right) \in L^2\left(0,T;L^2(\Omega)\right)$. ■

Remark 8.272 (On the Stability Estimate) The stability bound in (8.315) does not depend on ν^{-1} like for the Galerkin finite element discretization, see (8.303), but on a viscosity term that is on no account smaller. □

Theorem 8.273 (Error Estimate for the Velocity with Constants Depending on a Modified Viscosity) *Let $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, be a bounded domain with polyhedral and Lipschitz continuous boundary, let the regularity assumptions* (8.311) *and* (8.312) *be fulfilled and assume for the turbulent viscosity that*

$$
\nu_{\mathrm{T}} \in L^\infty\left(0,T;L^\infty(\Omega)\right). \tag{8.317}
$$

Let $(\boldsymbol{u},p) \in V \times Q$ be the solution of (8.302)*, let $V^h \subset V$ and $Q^h \subset Q$ be finite element spaces that satisfy the discrete inf-sup condition* (3.51)*, and let $\boldsymbol{w}^h \in V^h$ be the velocity solution of the three-scale projection-based VMS method* (8.310)*. Concerning the projection-based VMS method, it is assumed that $L^H \subseteq \mathbb{D}\left(V^h\right)$ (otherwise the eddy viscosity model would be subtracted from scales where it was not added previously). Then, the error $\boldsymbol{u} - \boldsymbol{w}^h$ satisfies for $t \in (0,T]$*

$$
\begin{aligned}
&\left\|(\boldsymbol{u}-\boldsymbol{w}^h)(t)\right\|_{L^2(\Omega)}^2 + \int_0^t \sum_{K\in\mathcal{T}^h} \nu_{\mathrm{mod}}^K\left(\boldsymbol{w}^h,\boldsymbol{w}^h - I_{\mathrm{St}}^h\boldsymbol{u}\right) \left\|\mathbb{D}\left(\boldsymbol{u}-\boldsymbol{w}^h\right)\right\|_{L^2(K)}^2 d\tau \\
&\le C\Bigg\{ \left\|(\boldsymbol{u}-I_{\mathrm{St}}^h\boldsymbol{u})(t)\right\|_{L^2(\Omega)}^2 \\
&\quad + \int_0^t \sum_{K\in\mathcal{T}^h} \nu_{\mathrm{mod}}^K\left(\boldsymbol{w}^h,\boldsymbol{w}^h - I_{\mathrm{St}}^h\boldsymbol{u}\right) \left\|\mathbb{D}\left(\boldsymbol{u}-I_{\mathrm{St}}^h\boldsymbol{u}\right)\right\|_{L^2(K)}^2\, d\tau
\end{aligned}
$$

$$
\begin{aligned}
&+\exp\left(\int_0^t \frac{C}{\left(\nu_{\mathrm{mod}}^{\min}\left(\boldsymbol{w}^h,\boldsymbol{w}^h-I_{\mathrm{St}}^h\boldsymbol{u}\right)\right)^3}\left\|\mathbb{D}\left(\boldsymbol{u}\right)\right\|_{L^2(\Omega)}^4\right)\\
&\times\Bigg[\left\|\boldsymbol{w}_0^h-I_{\mathrm{St}}^h\boldsymbol{u}(0)\right\|_{L^2(\Omega)}^2+\int_0^t\sum_{K\in\mathcal{T}^h}\nu_{\mathrm{mod}}^K\left(\boldsymbol{w}^h,\boldsymbol{u}-I_{\mathrm{St}}^h\boldsymbol{u}\right)\left\|\mathbb{D}\left(\boldsymbol{u}-I_{\mathrm{St}}^h\boldsymbol{u}\right)\right\|_{L^2(K)}^2\,d\tau\\
&+\int_0^t\frac{1}{\nu_{\mathrm{mod}}^{\min}\left(\boldsymbol{w}^h,\boldsymbol{w}^h-I_{\mathrm{St}}^h\boldsymbol{u}\right)}\Bigg(\left\|\partial_t\left(\boldsymbol{u}-I_{\mathrm{St}}^h\boldsymbol{u}\right)\right\|_{H^{-1}(\Omega)}^2\\
&+\left\|\mathbb{D}\left(\boldsymbol{u}-I_{\mathrm{St}}^h\boldsymbol{u}\right)\right\|_{L^2(\Omega)}^2\left\|\mathbb{D}\left(\boldsymbol{u}\right)\right\|_{L^2(\Omega)}^2+\inf_{q^h\in L^2(0,t;Q^h)}\left\|p-q^h\right\|_{L^2(\Omega)}^2\Bigg)\,d\tau\\
&+\frac{1}{\min_{\tau\in(0,t]}\left(\nu_{\mathrm{mod}}^{\min}\left(\boldsymbol{w}^h,\boldsymbol{w}^h\right)\right)^{1/2}}\\
&\times\left(\left\|\boldsymbol{w}^h(0)\right\|_{L^2(\Omega)}^2+\int_0^t\frac{\|\boldsymbol{f}\|_{H^{-1}(\Omega)}^2}{\nu_{\mathrm{mod}}^{\min}\left(\boldsymbol{w}^h,\boldsymbol{w}^h\right)}\,d\tau\right)\left\|\mathbb{D}\left(\boldsymbol{u}-I_{\mathrm{St}}^h\boldsymbol{u}\right)\right\|_{L^4(0,t;L^2)}^2\\
&+\int_0^t\sum_{K\in\mathcal{T}^h}\nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right)\left\|(I-P_{L^H})\mathbb{D}\left(\boldsymbol{u}\right)\right\|_{L^2(K)}^2\,d\tau\Bigg]\Bigg\},
\end{aligned}
\tag{8.318}
$$

where $\nu_{\mathrm{mod}}^{\min}$ *is defined in* (8.314) *and* $I_{\mathrm{St}}^h\boldsymbol{u}$ *is the Stokes projection of* $\boldsymbol{u}$ *defined in* (4.54)*, for which it is assumed that*

$$
\partial_t I_{\mathrm{St}}^h\boldsymbol{u}\in L^2\left(0,T;V'\right).
\tag{8.319}
$$

Proof The proof of the finite element error estimate follows the lines of the proof of the estimate for the Galerkin discretization, compare Theorem 7.35. For the three-scale projection-based VMS method, one obtains on the left-hand side of the error equation a modified viscous term and on the right-hand side two additional terms compared with the Galerkin discretization. Then, one estimates all terms on the right-hand side with the modified viscosity.

1. *Derivation of an error equation and splitting of the error.* The same splitting (7.46) as for the Galerkin finite element method is used

$$
\boldsymbol{e}(t)=\boldsymbol{u}(t)-\boldsymbol{w}^h(t)=\left(\boldsymbol{u}(t)-I_{\mathrm{St}}^h\boldsymbol{u}(t)\right)+\left(I_{\mathrm{St}}^h\boldsymbol{u}(t)-\boldsymbol{w}^h(t)\right)=\boldsymbol{\eta}(t)-\boldsymbol{\phi}^h(t).
$$

With the same arguments as in the proof of Theorem 7.35, one finds that

$$\mathbb{D}\left(I_{\mathrm{St}}^h \boldsymbol{u}\right) \in L^4\left(0, T; L^2(\Omega)\right). \tag{8.320}$$

Subtracting the discrete equation (8.310) from the continuous Eq. (8.302) gives for the test function $\boldsymbol{\phi}^h$ with a straightforward calculation the error equation

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 + \sum_{K\in\mathcal{T}^h} \nu_{\mathrm{mod}}^K\left(\boldsymbol{w}^h, \boldsymbol{\phi}^h\right)\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(K)}^2 \\
&= (\partial_t\boldsymbol{\eta}, \boldsymbol{\phi}^h) + (2\nu\mathbb{D}(\boldsymbol{\eta}), \mathbb{D}\left(\boldsymbol{\phi}^h\right)) + +n_{\mathrm{skew}}(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{\phi}^h) - n_{\mathrm{skew}}(\boldsymbol{w}^h, \boldsymbol{w}^h, \boldsymbol{\phi}^h) \\
&\quad - \left(\nabla\cdot\boldsymbol{\phi}^h, p - q^h\right) + \sum_{K\in\mathcal{T}^h} \nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right)\left((I - P_{L^H})\mathbb{D}(\boldsymbol{\eta}), (I - P_{L^H})\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right)_K \\
&\quad \sum_{K\in\mathcal{T}^h} \nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right)\left((I - P_{L^H})\mathbb{D}(\boldsymbol{u}), (I - P_{L^H})\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right)_K
\end{aligned} \tag{8.321}$$

with arbitrary $q^h \in Q^h$.

2. *Estimate all terms on the right hand-side of the error Eq.* (8.321). All bilinear terms in (8.321) are estimated more or less in the same way: using the Cauchy-Schwarz inequality (or the estimate for the dual pairing), Korn's inequality (3.43), and Young's inequality (A.5). The treatment of the local viscosity terms is done analogously as it was presented in detail in estimate (8.316). One obtains for the term with the temporal derivative

$$\begin{aligned}
\left(\partial_t\boldsymbol{\eta}, \boldsymbol{\phi}^h\right) &\le \|\partial_t\boldsymbol{\eta}\|_{H^{-1}(\Omega)}\left\|\nabla\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \le C\|\partial_t\boldsymbol{\eta}\|_{H^{-1}(\Omega)}\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)} \\
&\le \frac{C}{\nu_{\mathrm{mod}}^{\min}\left(\boldsymbol{w}^h, \boldsymbol{\phi}^h\right)}\|\partial_t\boldsymbol{\eta}\|_{H^{-1}(\Omega)}^2 + \frac{1}{8}\sum_{K\in\mathcal{T}^h} \nu_{\mathrm{mod}}^K\left(\boldsymbol{w}^h, \boldsymbol{\phi}^h\right)\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(K)}^2,
\end{aligned}$$

the viscous term

$$\begin{aligned}
\left(2\nu\mathbb{D}(\boldsymbol{\eta}), \mathbb{D}\left(\boldsymbol{\phi}^h\right)\right) &\le 2\nu\|\mathbb{D}(\boldsymbol{\eta})\|_{L^2(\Omega)}\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)} \\
&\le 8\nu\|\mathbb{D}(\boldsymbol{\eta})\|_{L^2(\Omega)}^2 + \frac{\nu}{8}\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)}^2,
\end{aligned}$$

the term with the pressure, where in addition (3.170) is used,

$$\begin{aligned}
&(\nabla\cdot\boldsymbol{\phi}^h, p - q^h) \\
&\le \left\|p - q^h\right\|_{L^2(\Omega)}\left\|\nabla\cdot\boldsymbol{\phi}^h\right\|_{L^2(\Omega)} \le C\left\|p - q^h\right\|_{L^2(\Omega)}\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)} \\
&\le \frac{C}{\nu_{\mathrm{mod}}^{\min}\left(\boldsymbol{w}^h, \boldsymbol{\phi}^h\right)}\left\|p - q^h\right\|_{L^2(\Omega)}^2 + \frac{1}{8}\sum_{K\in\mathcal{T}^h} \nu_{\mathrm{mod}}^K\left(\boldsymbol{w}^h, \boldsymbol{\phi}^h\right)\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(K)}^2,
\end{aligned}$$

the first additional term, using the definition (8.306),

$$\begin{aligned}
&\sum_{K\in\mathcal{T}^h} \nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right)\left((I-P_{L^H})\mathbb{D}\left(\boldsymbol{\eta}\right),(I-P_{L^H})\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right)_K \\
&\le 4\sum_{K\in\mathcal{T}^h} \nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right)\|(I-P_{L^H})\mathbb{D}\left(\boldsymbol{\eta}\right)\|_{L^2(K)}^2 \\
&\quad +\frac{1}{16}\sum_{K\in\mathcal{T}^h} \nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right)\left\|(I-P_{L^H})\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(K)}^2 \\
&= 4\sum_{K\in\mathcal{T}^h} \nu_{\mathrm{VMS}}^K\left(\boldsymbol{w}^h,\boldsymbol{\eta}\right)\|\mathbb{D}\left(\boldsymbol{\eta}\right)\|_{L^2(K)}^2+\frac{1}{16}\sum_{K\in\mathcal{T}^h} \nu_{\mathrm{VMS}}^K\left(\boldsymbol{w}^h,\boldsymbol{\phi}^h\right)\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(K)}^2,
\end{aligned}$$

and in a similar way for the second additional term

$$\begin{aligned}
&\sum_{K\in\mathcal{T}^h} \nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right)\left((I-P_{L^H})\mathbb{D}\left(\boldsymbol{u}\right),(I-P_{L^H})\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right)_K \\
&\le 4\sum_{K\in\mathcal{T}^h} \nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right)\|(I-P_{L^H})\mathbb{D}\left(\boldsymbol{u}\right)\|_{L^2(K)}^2+\frac{1}{16}\sum_{K\in\mathcal{T}^h} \nu_{\mathrm{VMS}}^K\left(\boldsymbol{w}^h,\boldsymbol{\phi}^h\right)\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(K)}^2.
\end{aligned}$$

The trilinear terms are decomposed as in (6.65), leading to

$$\begin{aligned}
&\left|n_{\mathrm{skew}}\left(\boldsymbol{u},\boldsymbol{u},\boldsymbol{\phi}^h\right)-n_{\mathrm{skew}}\left(\boldsymbol{w}^h,\boldsymbol{w}^h,\boldsymbol{\phi}^h\right)\right| \\
&\le\left|n_{\mathrm{skew}}\left(\boldsymbol{\eta},\boldsymbol{u},\boldsymbol{\phi}^h\right)\right|+\left|n_{\mathrm{skew}}\left(\boldsymbol{\phi}^h,\boldsymbol{u},\boldsymbol{\phi}^h\right)\right|+\left|n_{\mathrm{skew}}\left(\boldsymbol{w}^h,\boldsymbol{\eta},\boldsymbol{\phi}^h\right)\right|.
\end{aligned}$$

Applying (6.41) with $s=1/2$ and Young's inequality yields

$$\begin{aligned}
n_{\mathrm{skew}}(\boldsymbol{\eta},\boldsymbol{u},\boldsymbol{\phi}^h) &\le C\,\|\boldsymbol{\eta}\|_{L^2(\Omega)}^{1/2}\|\mathbb{D}\left(\boldsymbol{\eta}\right)\|_{L^2(\Omega)}^{1/2}\|\mathbb{D}\left(\boldsymbol{u}\right)\|_{L^2(\Omega)}\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)} \\
&\le\frac{C}{\nu_{\mathrm{mod}}^{\min}\left(\boldsymbol{w}^h,\boldsymbol{\phi}^h\right)}\|\boldsymbol{\eta}\|_{L^2(\Omega)}\|\mathbb{D}\left(\boldsymbol{\eta}\right)\|_{L^2(\Omega)}\|\mathbb{D}\left(\boldsymbol{u}\right)\|_{L^2(\Omega)}^2 \\
&\quad+\frac{1}{8}\sum_{K\in\mathcal{T}^h}\nu_{\mathrm{mod}}^K\left(\boldsymbol{w}^h,\boldsymbol{\phi}^h\right)\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(K)}^2,
\end{aligned}$$

$$\begin{aligned}
n_{\mathrm{skew}}(\boldsymbol{w}^h,\boldsymbol{\eta},\boldsymbol{\phi}^h) &\le C\left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)}^{1/2}\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^2(\Omega)}^{1/2}\|\mathbb{D}\left(\boldsymbol{\eta}\right)\|_{L^2(\Omega)}\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)} \\
&\le+\frac{C}{\nu_{\mathrm{mod}}^{\min}\left(\boldsymbol{w}^h,\boldsymbol{\phi}^h\right)}\left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)}\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^2(\Omega)}\|\mathbb{D}\left(\boldsymbol{\eta}\right)\|_{L^2(\Omega)}^2 \\
&\quad+\frac{1}{8}\sum_{K\in\mathcal{T}^h}\nu_{\mathrm{mod}}^K\left(\boldsymbol{w}^h,\boldsymbol{\phi}^h\right)\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(K)}^2,
\end{aligned}$$

and

$$\begin{aligned} n_{\mathrm{skew}}(\boldsymbol{\phi}^h, \boldsymbol{u}, \boldsymbol{\phi}^h) &\le C \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^{1/2} \|\mathbb{D}(\boldsymbol{u})\|_{L^2(\Omega)} \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(\Omega)}^{3/2} \\ &\le \frac{C}{\left(\nu_{\mathrm{mod}}^{\min}\left(\boldsymbol{w}^h, \boldsymbol{\phi}^h\right)\right)^3} \left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 \|\mathbb{D}(\boldsymbol{u})\|_{L^2(\Omega)}^4 \\ &\quad + \frac{1}{8} \sum_{K\in\mathcal{T}^h} \nu_{\mathrm{mod}}^K\left(\boldsymbol{w}^h, \boldsymbol{\phi}^h\right) \left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(K)}^2 . \end{aligned}$$

Collecting terms and using the definition (8.313) of ν_{mod}^K gives

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2 + \frac{1}{4}\sum_{K\in\mathcal{T}^h} \nu_{\mathrm{mod}}^K\left(\boldsymbol{w}^h, \boldsymbol{\phi}^h\right)\left\|\mathbb{D}\left(\boldsymbol{\phi}^h\right)\right\|_{L^2(K)}^2 \\ &\le C\left\{ \frac{1}{\nu_{\mathrm{mod}}^{\min}\left(\boldsymbol{w}^h, \boldsymbol{\phi}^h\right)}\Big[\|\partial_t\boldsymbol{\eta}\|_{H^{-1}(\Omega)}^2 + \|\boldsymbol{\eta}\|_{L^2(\Omega)}\|\mathbb{D}(\boldsymbol{\eta})\|_{L^2(\Omega)}\|\mathbb{D}(\boldsymbol{u})\|_{L^2(\Omega)}^2 \right. \\ &\quad + \left\|\boldsymbol{w}^h\right\|_{L^2(\Omega)}\left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^2(\Omega)}\|\mathbb{D}(\boldsymbol{\eta})\|_{L^2(\Omega)}^2 + \left\|p - q^h\right\|_{L^2(\Omega)}^2 \Big] \\ &\quad + \sum_{K\in\mathcal{T}^h}\nu_{\mathrm{mod}}^K\left(\boldsymbol{w}^h, \boldsymbol{\eta}\right)\|\mathbb{D}(\boldsymbol{\eta})\|_{L^2(K)}^2 + \sum_{K\in\mathcal{T}^h}\nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right)\|(I - P_{L^H})\mathbb{D}(\boldsymbol{u})\|_{L^2(K)}^2 \\ &\quad \left. + \frac{C}{\left(\nu_{\mathrm{mod}}^{\min}\left(\boldsymbol{w}^h, \boldsymbol{\phi}^h\right)\right)^3}\left\|\boldsymbol{\phi}^h\right\|_{L^2(\Omega)}^2\|\mathbb{D}(\boldsymbol{u})\|_{L^2(\Omega)}^4 \right\}. \end{aligned} \tag{8.322}$$

3. *Application of Gronwall's lemma A.54.* For applying Gronwall's lemma, the $L^1(0,T)$ regularity of the terms appearing in (8.322) has to be proved. From assumption (8.317), it follows with (8.317), (8.313), and (8.314) that

$$\nu_{\mathrm{mod}}^K\left(\boldsymbol{w}^h, \boldsymbol{\eta}\right) < \infty, \quad \nu_{\mathrm{VMS}}^K\left(\boldsymbol{w}^h, \boldsymbol{u}\right) < \infty, \quad \frac{1}{\nu_{\mathrm{mod}}^{\min}\left(\boldsymbol{w}^h, \boldsymbol{\phi}^h\right)} \in L^\infty\left(0, T; L^\infty(\Omega)\right).$$

Thus, the viscosity terms on the right-hand side of (8.322) are bounded in Ω for all times and it suffices to consider the norms of the functions. One obtains in the same way as in (7.54), with the gradient replaced by the deformation tensor,

$$\|\boldsymbol{\eta}\|_{L^2(\Omega)}\|\mathbb{D}(\boldsymbol{\eta})\|_{L^2(\Omega)}\|\mathbb{D}(\boldsymbol{u})\|_{L^2(\Omega)}^2 \le C\|\mathbb{D}(\boldsymbol{\eta})\|_{L^4(0,t;L^2(\Omega))}^2\|\mathbb{D}(\boldsymbol{u})\|_{L^4(0,t;L^2(\Omega))}^2 < \infty,$$

because of the regularity assumptions (8.312) and (8.320). The estimate of the other term is performed in the same way as estimate (7.55), but now the stability estimate (8.315) is applied

$$
\begin{aligned}
&\int_0^t \left\|\boldsymbol{w}^h(\tau)\right\|_{L^2(\Omega)} \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)(\tau)\right\|_{L^2(\Omega)} \left\|\mathbb{D}\left(\boldsymbol{\eta}\right)(\tau)\right\|_{L^2(\Omega)}^2 d\tau \\
&\le \left\|\boldsymbol{w}^h\right\|_{L^\infty(0,t;L^2)} \left\|\mathbb{D}\left(\boldsymbol{w}^h\right)\right\|_{L^2(0,t;L^2)} \left\|\mathbb{D}\left(\boldsymbol{\eta}\right)\right\|_{L^4(0,t;L^2)}^2 \\
&\le \frac{C}{\min_{\tau\in(0,t]}\left(\nu_{\mathrm{mod}}^{\min}\left(\boldsymbol{w}^h,\boldsymbol{w}^h\right)\right)^{1/2}} \\
&\quad \times \left(\left\|\boldsymbol{w}^h(0)\right\|_{L^2(\Omega)}^2 + \int_0^t \frac{\|\boldsymbol{f}\|_{H^{-1}(\Omega)}^2}{\nu_{\mathrm{mod}}^{\min}\left(\boldsymbol{w}^h,\boldsymbol{w}^h\right)}\, d\tau\right) \left\|\mathbb{D}\left(\boldsymbol{\eta}\right)\right\|_{L^4(0,t;L^2)}^2 < \infty.
\end{aligned}
$$

The $L^1(0,T)$-regularity of the other terms is a direct consequence of (8.312), the stability of the Stokes projection (4.55), (8.319), and (8.320). Now, Gronwall's inequality can be applied.

4. *Application of the triangle inequality.* The application of the triangle inequality is done in the same way as in the proof of Theorem 7.35. ■

Remark 8.274 (On Error Estimate (8.318)*)*

- Even if the constants in the error bound (8.318) depend on a modified viscosity that is on no account smaller than ν, the bound itself depends on inverse powers of ν via the norms of $\boldsymbol{u}$, e.g., see the stability estimate (8.303) for the Navier–Stokes equations. A dependency of this kind is inevitable if the error to the solution of the Navier–Stokes equations is studied.
- All terms in the error bound (8.318) but the last one contain factors with interpolation errors. The last term tends to zero if $\nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right) \to 0$ or if L^H tends to $\mathbb{D}(V)$. In both situations, the Galerkin finite element discretization of the Navier–Stokes equations is recovered asymptotically. Otherwise, in particular if $\nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right)$ and L^H are fixed and $h \to 0$, one cannot expect that the solution of the discrete system with the additional viscosity term tends to the solution of the continuous Navier–Stokes equations where such a term is not present.

  Note that in practice $\nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right)$ depends on the (local) mesh width, e.g., if Smagorinsky-type models of form (8.290)–(8.292) are used then generally $\delta = \mathcal{O}(h)$.
- For fixed h and $\nu_{\mathrm{T}}^K\left(\boldsymbol{w}^h\right) \to 0$, the error bound (8.318) tends to the error bound (7.44) for the Galerkin discretization of the Navier–Stokes equations (with the gradient replaced by the deformation tensor).
- There is no improvement in the asymptotic of the exponential if

$$\nu_{\mathrm{mod}}^{\min}\left(\boldsymbol{w}^h, \boldsymbol{w}^h - I_{\mathrm{St}}^h\boldsymbol{u}\right) = 2\nu$$

  or equivalently if

$$\nu_{\mathrm{VMS}}^K\left(\boldsymbol{w}^h, \boldsymbol{w}^h - I_{\mathrm{St}}^h\boldsymbol{u}\right) = 0,$$

i.e.,

$$\left\|P_{L^H}\mathbb{D}\left(\boldsymbol{w}^h - I_{\mathrm{St}}^h\boldsymbol{u}\right)\right\|_{L^2(K)} = \left\|\mathbb{D}\left(\boldsymbol{w}^h - I_{\mathrm{St}}^h\boldsymbol{u}\right)\right\|_{L^2(K)}$$

for all $K \in \mathcal{T}^h$ and all times. By squaring and summing up, it follows that $\left\|P_{L^H}\mathbb{D}\left(\boldsymbol{w}^h - I_{\mathrm{St}}^h\boldsymbol{u}\right)\right\|_{L^2(\Omega)}^2 = \left\|\mathbb{D}\left(\boldsymbol{w}^h - I_{\mathrm{St}}^h\boldsymbol{u}\right)\right\|_{L^2(\Omega)}^2$. From the Pythagorean theorem, compare (8.305),

$$\begin{aligned}&\left\|\mathbb{D}\left(\boldsymbol{w}^h - I_{\mathrm{St}}^h\boldsymbol{u}\right)\right\|_{L^2(\Omega)}^2\\ &= \left\|P_{L^H}\mathbb{D}\left(\boldsymbol{w}^h - I_{\mathrm{St}}^h\boldsymbol{u}\right)\right\|_{L^2(\Omega)}^2 + \left\|\left(I - P_{L^H}\right)\mathbb{D}\left(\boldsymbol{w}^h - I_{\mathrm{St}}^h\boldsymbol{u}\right)\right\|_{L^2(\Omega)}^2,\end{aligned}$$

it follows that

$$\left(I - P_{L^H}\right)\mathbb{D}\left(\boldsymbol{w}^h\right) = \left(I - P_{L^H}\right)\mathbb{D}\left(I_{\mathrm{St}}^h\boldsymbol{u}\right),$$

i.e., the small resolved scales of the discrete velocity solution and the Stokes projection are the same. This situation is unlikely for turbulent flows since the small resolved scales of $\boldsymbol{w}^h$ are considerably influenced by the eddy viscosity model whereas the Stokes projection does not possess any information about this model Hence, this situation is only likely if there are no small resolved scales in the flow but only large scales, which is not the case in turbulent flows.

□

8.8.6.4 Implementation and Numerical Experience

Remark 8.275 (Implementation of Method (8.288)) The version of the three-scale projection-based VMS method that uses all spaces on the same grid $\mathcal{T}^h$, see Remark 8.261, can be implemented efficiently if the space L^H possesses the following two properties:

- the space L^H is a discontinuous finite element space with respect to $\mathcal{T}^h$,
- the basis of L^H is $L^2(\Omega)$ orthogonal.

Here, some details of the implementation will be described.

Let the velocity vector $\boldsymbol{w}^h$ and the symmetric tensor $\mathbb{G}^H$ be given by

$$\boldsymbol{w}^h = \begin{pmatrix} w_1^h \\ w_2^h \\ w_3^h \end{pmatrix}, \quad \mathbb{G}^H = \begin{pmatrix} \underline{g_{11}} & \underline{g_{12}} & \underline{g_{13}} \\ \underline{g_{12}} & \underline{g_{22}} & \underline{g_{23}} \\ \underline{g_{13}} & \underline{g_{23}} & \underline{g_{33}} \end{pmatrix}$$

and let the spaces V^h and L^H be equipped with the bases

$$V^h = \operatorname{span}\left\{ \left\{ \begin{pmatrix} \phi_i^h \\ 0 \\ 0 \end{pmatrix} \right\}_{i=1}^{N_v} \cup \left\{ \begin{pmatrix} 0 \\ \phi_i^h \\ 0 \end{pmatrix} \right\}_{i=1}^{N_v} \cup \left\{ \begin{pmatrix} 0 \\ 0 \\ \phi_i^h \end{pmatrix} \right\}_{i=1}^{N_v} \right\},$$

$$L^H = \operatorname{span}\left\{ \left\{ \begin{pmatrix} l_j^H & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\}_{j=1}^{N_L}, \left\{ \frac{1}{2} \begin{pmatrix} 0 & l_j^H & 0 \\ l_j^H & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\}_{j=1}^{N_L}, \left\{ \frac{1}{2} \begin{pmatrix} 0 & 0 & l_j^H \\ 0 & 0 & 0 \\ l_j^H & 0 & 0 \end{pmatrix} \right\}_{j=1}^{N_L}, \right.$$

$$\left. \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & l_j^H & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\}_{j=1}^{N_L}, \left\{ \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & l_j^H \\ 0 & l_j^H & 0 \end{pmatrix} \right\}_{j=1}^{N_L}, \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & l_j^H \end{pmatrix} \right\}_{j=1}^{N_L} \right\}.$$

After a discretization of (8.288) in time and a linearization of the convective term in the current discrete time, one obtains a linear saddle point problem of the following form

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} & B_1^T & \tilde{G}_{11} & \tilde{G}_{12} & \tilde{G}_{13} & \tilde{G}_{14} & \tilde{G}_{15} & \tilde{G}_{16} \\ A_{21} & A_{22} & A_{23} & B_2^T & \tilde{G}_{21} & \tilde{G}_{22} & \tilde{G}_{23} & \tilde{G}_{24} & \tilde{G}_{25} & \tilde{G}_{26} \\ A_{31} & A_{32} & A_{33} & B_3^T & \tilde{G}_{31} & \tilde{G}_{32} & \tilde{G}_{33} & \tilde{G}_{34} & \tilde{G}_{35} & \tilde{G}_{36} \\ B_1 & B_2 & B_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ G_{11} & G_{12} & G_{13} & 0 & M & 0 & 0 & 0 & 0 & 0 \\ G_{21} & G_{22} & G_{23} & 0 & 0 & \frac{M}{2} & 0 & 0 & 0 & 0 \\ G_{31} & G_{32} & G_{33} & 0 & 0 & 0 & \frac{M}{2} & 0 & 0 & 0 \\ G_{41} & G_{42} & G_{43} & 0 & 0 & 0 & 0 & M & 0 & 0 \\ G_{51} & G_{52} & G_{53} & 0 & 0 & 0 & 0 & 0 & \frac{M}{2} & 0 \\ G_{61} & G_{62} & G_{63} & 0 & 0 & 0 & 0 & 0 & 0 & M \end{pmatrix} \begin{pmatrix} \underline{w_1} \\ \underline{w_2} \\ \underline{w_3} \\ \underline{p} \\ \underline{g_{11}} \\ \underline{g_{12}} \\ \underline{g_{13}} \\ \underline{g_{22}} \\ \underline{g_{23}} \\ \underline{g_{33}} \end{pmatrix} = \begin{pmatrix} \underline{f_1} \\ \underline{f_2} \\ \underline{f_3} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{8.323}$$

The matrices $A_{11}, \ldots, A_{33}$ and B_1, B_2, B_3 have to be assembled if (8.288) is discretized without the terms involving $\mathbb{G}^H$, i.e., if the turbulence model is applied to all scales. The matrix M in (8.323) is the mass matrix of L^H, i.e., $(M)_{ij} = (l_j^H, l_i^H)$. The general entries of the matrices $G_{11}, \ldots, G_{63}$ and $\tilde{G}_{11}, \ldots, \tilde{G}_{36}$ can be computed using the bases of V^h and L^H. Straightforward calculations give

$$
\begin{aligned}
(G_{11})_{ij} &= \left(\begin{pmatrix} \partial_x\phi_j^h & \partial_y\phi_j^h/2 & \partial_z\phi_j^h/2 \\ \partial_y\phi_j^h/2 & 0 & 0 \\ \partial_z\phi_j^h/2 & 0 & 0 \end{pmatrix}, \begin{pmatrix} l_i^H & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right) \\
&= \left(\partial_x\phi_j^h, l_i^H\right), \qquad G_{22} = G_{33} = \tfrac{1}{2}G_{11}, \\
(G_{42})_{ij} &= \left(\partial_y\phi_j^h, l_i^H\right), \qquad G_{21} = G_{53} = \tfrac{1}{2}G_{42}, \\
(G_{63})_{ij} &= \left(\partial_z\phi_j^h, l_i^H\right), \qquad G_{31} = G_{52} = \tfrac{1}{2}G_{63}, \\
(\tilde{G}_{11})_{ij} &= \left(\nu_{\mathrm{T}} l_j^H, \partial_x\phi_i^h\right), \qquad \tilde{G}_{22} = \tilde{G}_{33} = \tfrac{1}{2}\tilde{G}_{11}, \\
(\tilde{G}_{24})_{ij} &= \left(\nu_{\mathrm{T}} l_j^H, \partial_y\phi_i^h\right), \qquad \tilde{G}_{12} = \tilde{G}_{35} = \tfrac{1}{2}\tilde{G}_{24}, \\
(\tilde{G}_{36})_{ij} &= \left(\nu_{\mathrm{T}} l_j^H, \partial_z\phi_i^h\right), \qquad \tilde{G}_{13} = \tilde{G}_{33} = \tfrac{1}{2}\tilde{G}_{36}.
\end{aligned}
\tag{8.324}
$$

All other blocks $G_{\alpha\beta}$ and $\tilde{G}_{\alpha\beta}$ vanish. Thus, one has to assemble only the entries of the blocks $G_{11}, G_{42}, G_{63}, \tilde{G}_{11}, \tilde{G}_{24}, \tilde{G}_{36}$. If the space L^H is static, the matrices G_{11}, G_{42}, G_{63}, and M have to be assembled only once since they are not time-dependent. The matrices $\tilde{G}_{11}, \tilde{G}_{24}, \tilde{G}_{36}$ are time-dependent if ν_{T} is time-dependent, e.g., if ν_{T} depends on the discrete solution or if the space L^H changes in time.

Now, (8.323) can be solved for $\underline{g_{11}}, \ldots, \underline{g_{33}}$. This step leads to a saddle point problem for $(\underline{w_1}, \underline{w_2}, \underline{w_3}, \underline{p})$ of the form

$$
\begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} & \tilde{A}_{13} & B_1^T \\ \tilde{A}_{21} & \tilde{A}_{22} & \tilde{A}_{23} & B_2^T \\ \tilde{A}_{31} & \tilde{A}_{32} & \tilde{A}_{33} & B_3^T \\ B_1 & B_2 & B_3 & 0 \end{pmatrix} \begin{pmatrix} \underline{w_1} \\ \underline{w_2} \\ \underline{w_3} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f_1} \\ \underline{f_2} \\ \underline{f_3} \\ 0 \end{pmatrix}
\tag{8.325}
$$

with

$$
\begin{aligned}
\tilde{A}_{11} &= A_{11} - \tilde{G}_{11}M^{-1}G_{11} - \frac{1}{2}\tilde{G}_{24}M^{-1}G_{42} - \frac{1}{2}\tilde{G}_{36}M^{-1}G_{63}, \\
\tilde{A}_{12} &= A_{12} - \frac{1}{2}\tilde{G}_{24}M^{-1}G_{11}, \\
\tilde{A}_{13} &= A_{13} - \frac{1}{2}\tilde{G}_{36}M^{-1}G_{11}, \\
&\ \vdots \\
\tilde{A}_{33} &= A_{33} - \tilde{G}_{36}M^{-1}G_{63} - \frac{1}{2}\tilde{G}_{11}M^{-1}G_{11} - \frac{1}{2}\tilde{G}_{24}M^{-1}G_{42}.
\end{aligned}
$$

Note that a different scaling of the basis functions of L^H leads to the same result since the scaling factor cancels.

The (i,j)-th entry of $\tilde{G}_{11}M^{-1}G_{11}$ has the form

$$\begin{aligned}
&\left(\tilde{G}_{11}M^{-1}G_{11}\right)_{ij} \\
&= \sum_{m,n=1}^{N_L} (\tilde{G}_{11})_{im}(M^{-1})_{mn}(G_{11})_{nj} \\
&= \sum_{m,n=1}^{N_L} (\nu_{\mathrm{T}} l_m^H, \partial_x \phi_i^h)(M^{-1})_{mn}(\partial_x \phi_j^h, l_n^H) \qquad (8.326)\\
&= \sum_{m,n=1}^{N_L} \left(\sum_{K\in\mathcal{T}^h} (\nu_{\mathrm{T}} l_m^H, \partial_x \phi_i^h)_K \right)(M^{-1})_{mn} \left(\sum_{K\in\mathcal{T}^h} (\partial_x \phi_j^h, l_n^H)_K \right).
\end{aligned}$$

This formula reveals that for an efficient implementation of the three-scale projection-based VMS method the two requirements on L^H given above have to be fulfilled:

- An efficient computation of (8.326) is possible if M is a diagonal matrix. This situation is given if and only if the basis functions of L^H are $L^2(\Omega)$ orthogonal. One obtains in this case

$$\left(\tilde{G}_{11}M^{-1}G_{11}\right)_{ij} = \sum_{m=1}^{N_L} \frac{\sum_{K\in\mathcal{T}^h} (\nu_{\mathrm{T}} l_m^H, \partial_x \phi_i^h)_K \sum_{K\in\mathcal{T}^h} (\partial_x \phi_j^h, l_m^H)_K}{\sum_{K\in\mathcal{T}^h} (l_m^H, l_m^H)_K}. \qquad (8.327)$$

- For an efficient implementation, the sparsity pattern of $\tilde{A}_{\alpha\beta}$ must not be larger than the sparsity pattern of $A_{\alpha\beta}$. Then, the entries coming from terms like (8.327) can be simply added to $A_{\alpha\beta}$. An entry $(A_{\alpha\beta})_{ij}$ generally does not vanish if the intersection of the support of v_i^h and the support of v_j^h is at least one mesh cell K. If the support of l_m^H is only one mesh cell, then the numerator in (8.327) may be only not equal to zero if this mesh cell belongs also to the support of v_i^h and to the support of v_j^h. In this case, the sparsity pattern of $\tilde{G}_{11}M^{-1}G_{11}$, and hence of $\tilde{A}_{11}$, will be the same as of A_{11}. The requirement on the support of l_m^H is fulfilled if L^H is a discontinuous finite element space.

Thus, L^H has to consist of piecewise symmetric tensors and it has to be equipped with a basis that is $L^2(\Omega)$ orthogonal, e.g., with a basis of Legendre polynomials, see Example B.54 for spaces with these properties.

In summary, the main extension of an existing finite element code for the incompressible Navier–Stokes equations consists in the assembling of the matrices $\tilde{A}_{ij}$, instead of A_{ij}, using formulas like (8.327). □

Remark 8.276 (Adaptive Choice of the Projection Space L^H) The basic intentions for choosing the projection space adaptively are to focus the application of the eddy viscosity model to sub-regions where the flow is highly turbulent and to switch off the turbulence model in sub-regions with laminar flow. In John and Kaya (2005, 2008), John and Roland (2007) and John and Kindl (2010a), numerical studies with the three-scale projection-based VMS method were performed with the static spaces $L^H = P_0$ and $L^H = P_1^{\text{disc}}$. It was concluded in John and Kindl (2010a) that the choice of the projection space has a much higher impact on the results than the choice of the eddy viscosity model, compare also Example 8.277.

In John and Kindl (2010b), the three-scale projection-based VMS method with adaptively chosen projection space was introduced. In this method, the projection space might change during the simulations, either after every time step or in larger time intervals. In strongly turbulent subregions, the projection space is chosen to be locally small and the smaller the turbulence intensity is, the larger the projection space becomes. The local turbulence intensity is estimated with the size of the local small resolved scales

$$\eta_K = \frac{\left\| \mathbb{G}^H - \mathbb{D}\left(\boldsymbol{w}^h\right) \| \right\|_{L^2(K)}}{\|1\|_{L^2(K)}} = \frac{\left\| \mathbb{G}^H - \mathbb{D}\left(\boldsymbol{w}^h\right) \| \right\|_{L^2(K)}}{|K|^{1/2}} \quad \forall\, K \in \mathcal{T}^h. \qquad (8.328)$$

Then, the local projection space is assigned in the following way. Given three non-negative constants $C_1 \le C_2 \le C_3$ and a reference value η, e.g., some mean value of $\{\eta_K\}_{K\in\mathcal{T}^h}$ in space or space-time, the space $L^H(K)$ was chosen as follows:

- for cells K with $\eta_K/\eta \le C_1$ set $L^H(K) = P_2^{\text{disc}}(K)$, $\nu_{\text{T}}(K) = 0$, i.e., the turbulence model is switched off,
- for cells K with $C_1 < \eta_K/\eta \le C_2$ set $L^H(K) = P_1(K)$,
- for cells K with $C_2 < \eta_K/\eta \le C_3$ set $L^H(K) = P_0(K)$,
- for cells K with $C_3 < \eta_K/\eta$ set $L^H(K) = \{\mathbb{O}\}$, i.e., the turbulence model is applied on K to all resolved scales.

Note that since L^H is a discontinuous finite element space, the assignment of different local spaces on different mesh cells is no difficulty.

Numerical studies in John and Kindl (2010b) showed in fact that with the estimator (8.328) the space L^H corresponds to the expectations. For the turbulent channel flow at $\text{Re}_\tau = 180$, see Example D.12, it was $L^H(K) = \{\mathbb{O}\}$ at the walls and the turbulence model was switched off in the center of the channel. In simulations of the turbulent flow around a cylinder, Example D.13, the space $L^H(K) = \{\mathbb{O}\}$ was chosen at the cylinder In regions far away from the cylinder, which are not in the wake, the turbulence model was switched off, compare Fig. 8.17 below. The studies in John and Kindl (2010b) were performed on hexahedral meshes with the pair of spaces Q_2/P_1^{disc}, see Remark 3.141. The extension of the adaptive projection-based VMS method to tetrahedral grids with the Bernardi–Raugel element $P_2^{\text{BR}}/P_1^{\text{disc}}$, see Example 3.140, was presented in John et al. (2010). □

Example 8.277 (Turbulent Channel Flow at $Re_\tau = 180$*)* The definition of this example is given in Example D.12 and the setup of the simulations was exactly as described in Example 8.128.

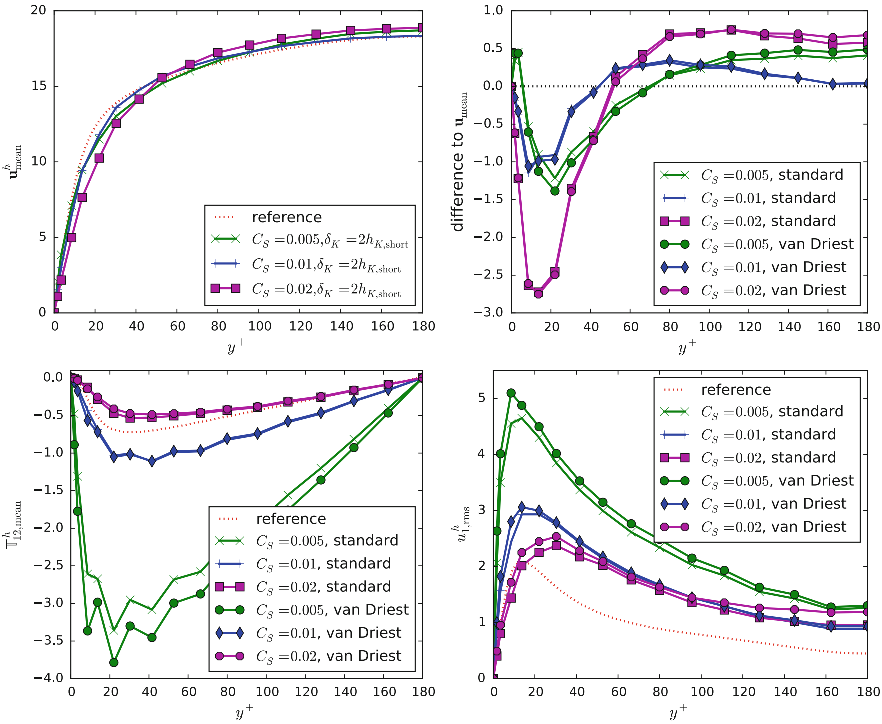

Fig. 8.11 Example 8.277. Statistics of interest obtained with the three-scale coarse space projection-based VMS method (8.288) with $L^H = P_0$

Results are presented in Figs. 8.11, 8.12 and 8.13 for the three-scale coarse space projection-based VMS method (8.288) with $L^H = P_0$, $L^H = P_1^{\text{disc}}$, and the adaptive choice of the projection space, respectively.

As already noted in Remark 8.216, a goal of three-scale VMS methods consists in reducing the influence of the Smagorinsky model by restricting its direct application to the small resolved scales. The achievement of this goal can be seen clearly by the fact that one gets reasonable results for a much wider range of parameters C_S than it was the case for the Smagorinsky LES model, compare with Fig. 8.4. Also the most accurate results for the VMS methods are obtained for larger values of C_S. As for the Smagorinsky LES model, there is only little impact of using the van Driest damping (8.152). Often, but not always, the results with the standard Smagorinsky method are a little bit more accurate. □

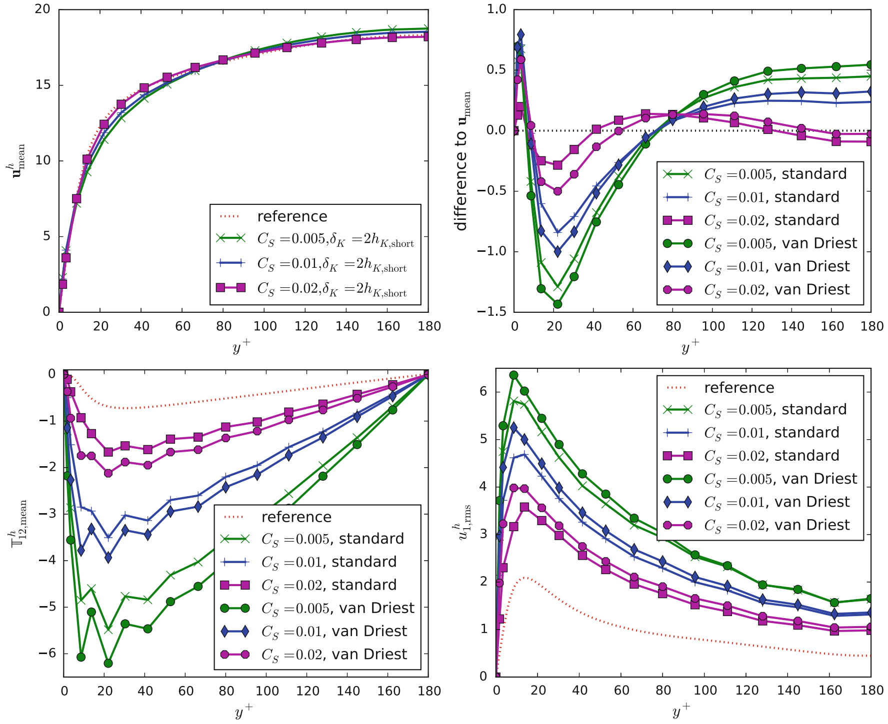

Fig. 8.12 Example 8.277. Statistics of interest obtained with the three-scale coarse space projection-based VMS method (8.288) with $L^H = P_1^{\mathrm{disc}}$

8.9 Comparison of Some Turbulence Models in Numerical Studies

Remark 8.278 (Contents) This section presents a few studies that compare the numerical solutions obtained with some of the turbulence models introduced in this chapter. It will be concentrated on the Smagorinsky model and the three-scale coarse space projection-based VMS method since these methods showed the most accurate results for the turbulent channel flow at $\mathrm{Re}_\tau = 180$. Simulations with the rational LES model, the Leray-α model, and the Navier–Stokes-α model for the problems presented in this section generally did not provide satisfactory solutions. $\square$

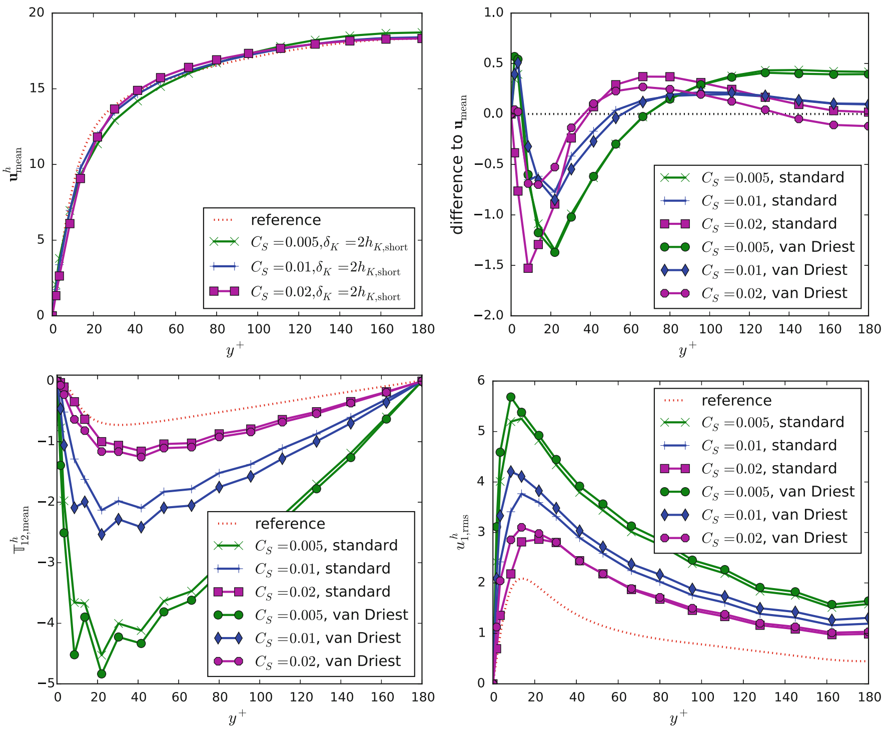

Fig. 8.13 Example 8.277. Statistics of interest obtained with the three-scale coarse space projection-based VMS method (8.288) with the adaptive choice of the projection space

Example 8.279 (Turbulent Channel Flow at $Re_\tau = 180$*)* In this example, only the best parameters for the considered turbulence models from Examples 8.128 and 8.277 will be presented. These parameters are:

- Smagorinsky model (8.69) (Smag): $\delta_K = 2h_{K,\text{short}}$ and $C_S = 0.005$, van Driest damping (8.152),
- coarse space projection-based method (8.288) with $L^h = P_0$ (VMS0): $\delta_K = 2h_{K,\text{short}}$ and $C_S = 0.01$, no van Driest damping,
- coarse space projection-based method (8.288) with $L^h = P_1^{\text{disc}}$ (VMS1): $\delta_K = 2h_{K,\text{short}}$ and $C_S = 0.02$, no van Driest damping,
- coarse space projection-based method (8.288) with adaptive choice of the projection space (VMSadap): $\delta_K = 2h_{K,\text{short}}$ and $C_S = 0.02$, van Driest damping (8.152).

The results for these parameters are collected in Fig. 8.14. It can be deduced that among the considered turbulence models, VMS0 and VMSadap gave the most accurate results and the Smagorinsky model the least accurate ones. □

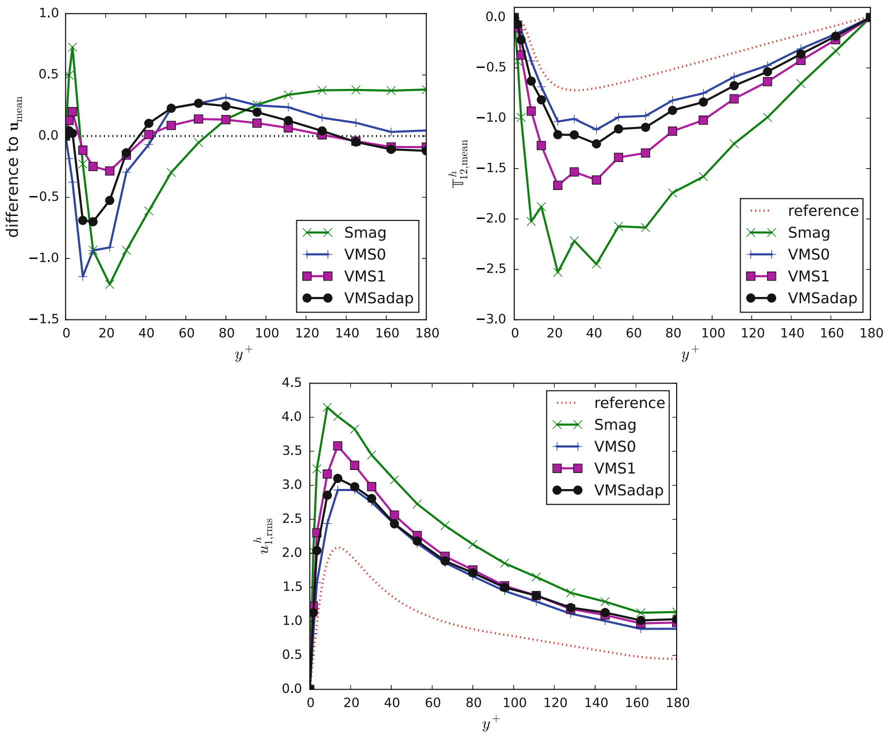

Fig. 8.14 Example 8.279. Statistics of interest for the best parameter choices of the Smagorinsky model and the three-scale coarse space projection-based VMS models for the turbulent channel flow at $\text{Re}_\tau = 180$

Example 8.280 (Turbulent Channel Flow at $Re_\tau = 395$) A description of this example can be found in Example D.12. The setup of the simulations was similar to the turbulent channel flow problem at $\text{Re}_\tau = 180$, compare Example 8.128. Only, finer grids in space and time were used. The spatial grid consisted of $16 \times 16 \times 32$ cells, leading to 199,680 velocity degrees of freedom (including Dirichlet nodes) and 32,768 pressure degrees of freedom for the Q_2/P_1^{disc} pair of finite element spaces. As time step in the Crank–Nicolson scheme, $\Delta t = 0.002$ was chosen. All other components of the setup, like the length of the time interval and the grading of the grid towards the walls, were the same as in Example 8.128.

Results are presented in Fig. 8.15. The best results were obtained with VMS0 and $C_S = 0.02$. Also some other results, like Smag with $C_S = 0.01$ and VMSadap with $C_S = 0.02$, are comparably good. Inaccurate results were computed using VMS1 with $C_S = 0.01$ and Smag with $C_S = 0.005$. Thus, the best parameter choice for the Smagorinsky model for the turbulent channel flow at $\text{Re}_\tau = 180$ was not an appropriate choice in this example. □

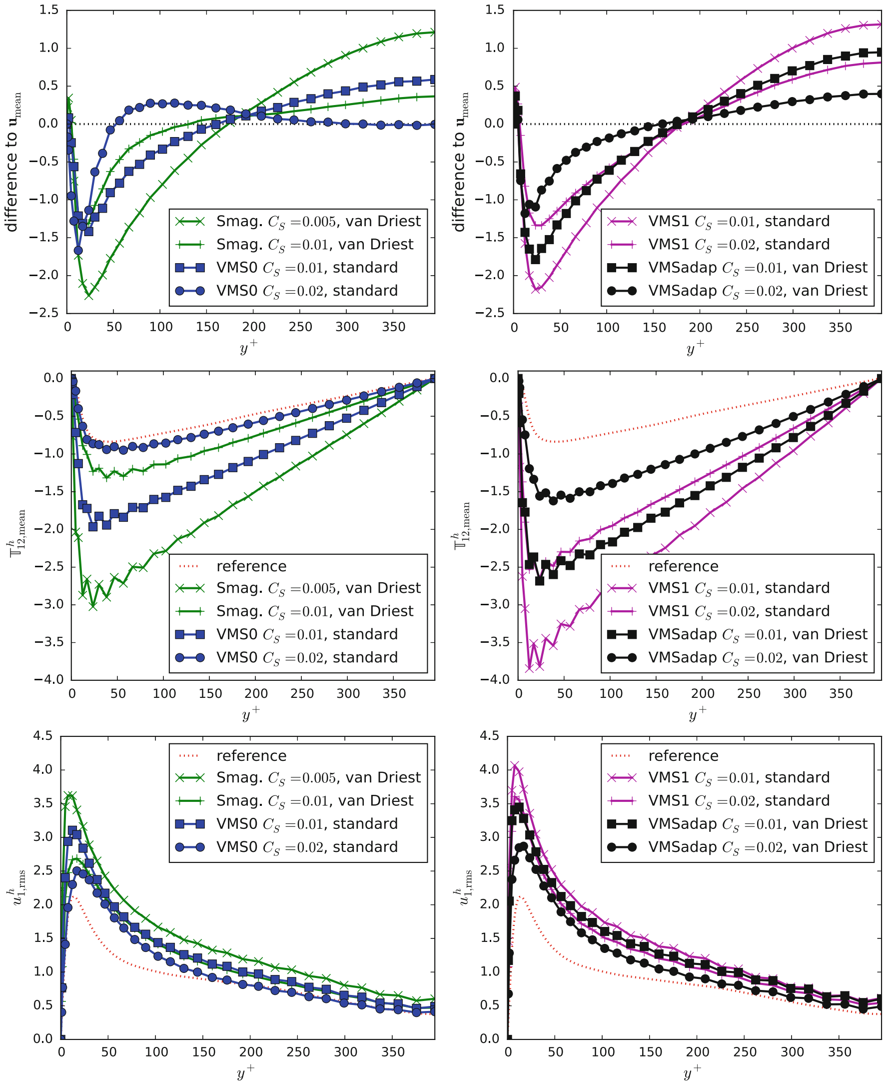

Fig. 8.15 Example 8.280. Statistics of interest for the Smagorinsky model and the three-scale coarse space projection-based VMS models for the turbulent channel flow at $\mathrm{Re}_\tau = 395$

Example 8.281 (Turbulent Flow Around a Cylinder at $Re = 22000$) This example and the quantities of interest are described in Example D.13. Simulations were performed on a hexahedral grid with the Q_2/P_1^{disc} pair of finite element spaces with 522,720 velocity degrees of freedom (including Dirichlet nodes) and 81,920 pressure degrees of freedom (level 2), see Fig. 8.16 for the initial grid (level 0). The Crank–Nicolson scheme was applied with the equidistant time step $\Delta t = 0.005$.

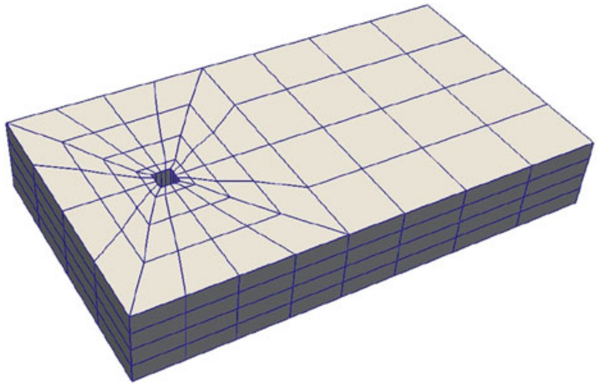

Fig. 8.16 Example 8.281. Initial grid, level 0

Table 8.1 Example 8.281

Method	C_S	$\overline{c}_{\text{lift}}$	$c_{\text{lift,rms}}$	$\overline{c}_{\text{drag}}$	$c_{\text{drag,rms}}$	St
Smag	0.005	Blow up				
Smag	0.01	0.001	1.46	2.58	0.11	0.151
Smag	0.02	0.023	1.52	2.52	0.16	0.140
VMS0	0.01	−0.017	1.28	2.60	0.19	0.139
VMS0	0.02	−0.015	1.29	2.65	0.14	0.142
VMS1	0.01	−0.004	1.04	2.49	0.15	0.136
VMS1	0.02	−0.026	1.00	2.50	0.21	0.131
VMSadap	0.01	−0.006	1.15	2.49	0.21	0.131
VMSadap	0.02	−0.026	1.49	2.55	0.16	0.141
Experimental values		0	[0.7, 1.4]	[1.9, 2.1]	[0.1, 0.2]	0.132

Computed values for the quantities of interest

Van Driest damping was not applied. All flows were allowed to develop for 10 s and then the statistics were computed for 30 periods.

Results computed with the considered methods are presented in Table 8.1. It can be observed that most of the obtained results are qualitatively similar: the root mean squared (rms) values are within the interval of the experimental values or only slightly outside and the mean drag coefficient is considerably larger than in the experiments. The latter observation can be found often in the literature, e.g., see Gravemeier et al. (2011) and Rasthofer and Gravemeier (2013). Only the Strouhal number shows some noticeable differences. It is worst with the Smagorinsky model with $C_S = 0.01$ and best with VMS1.

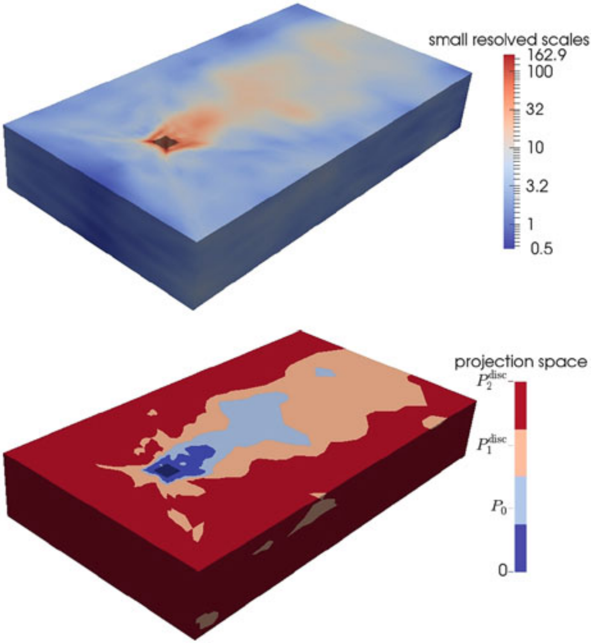

Fig. 8.17 Example 8.281. Snapshot of the size of the small resolved scales (*top*) and the corresponding projection space (*bottom*) in the three-level projection-based VMS method with adaptive projection space

In addition, it is remarkable that the Smagorinsky model with $C_S = 0.005$ blew up. Hence, the best parameter in the turbulent channel flow at $\text{Re}_\tau = 180$ failed in this example. In contrast, it can be seen also in this example that the three-scale coarse space projection-based VMS methods are much less sensitive to the choice of C_S.

Figure 8.17 provides an impression of the size and the distribution of the small resolved scales and the corresponding choice of the adaptive projection space. In this example, the highest turbulence intensity is expected at the cylinder and downstream the cylinder. It can be observed that the size of the small resolved scales corresponds well to this expectation. Thus, the quantity η_K from (8.328) is an appropriate estimate of the local turbulence intensity. In the picture of the projection space, it can be seen that the turbulence model has the most impact at the cylinder, the impact becomes smaller downstream the cylinder, and it is essentially switched off in the other regions. □

Example 8.282 (Turbulent Flow Around a Wall-Mounted Cube at Re $=$ 40000*)* This problem is described in Example D.14.

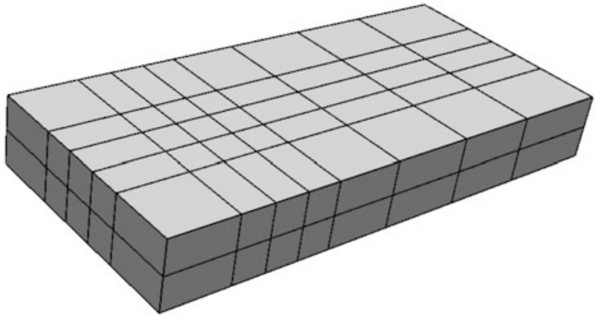

Fig. 8.18 Example 8.282. Initial grid, level 0

Table 8.2 Example 8.282

Level	Velocity	Pressure	All
2	134,739	20,224	154,963
3	1,023,651	161,792	1,185,443

Number of degrees of freedom in space (including Dirichlet nodes)

Table 8.3 Example 8.282

Method	C_S	Level 2	Level 3
Smag	0.005	1.16	1.46
Smag	0.01	1.26	1.45
VMS0	0.01	1.19	1.47
VMS0	0.02	1.23	1.45
VMS1	0.01	1.09	1.39
VMS1	0.02	1.11	1.38
VMSadap	0.01	1.13	1.37
VMSadap	0.02	1.21	1.33

Mean drag coefficients, reference $\overline{c}_{\mathrm{drag}} \approx 1.5$

Numerical simulations were performed with the Q_2/P_1^{disc} pair of finite element spaces on hexahedral grids. The initial grid (level 0) is depicted in Fig. 8.18 and information on the number of degrees of freedom are provided in Table 8.2. The time interval $[0, 10]$ was considered. As temporal discretization, the Crank–Nicolson scheme with equidistant time step $\Delta t = 0.01$ was applied. Averaging in time of the computed drag coefficients was performed in the interval $[5, 10]$. All turbulence models were applied without van Driest damping.

Results for the different turbulence models are presented in Table 8.3. The temporal evolution of the computed drag coefficient is exemplary shown for two models in Fig. 8.19. It can be seen that all turbulence models underpredict the mean

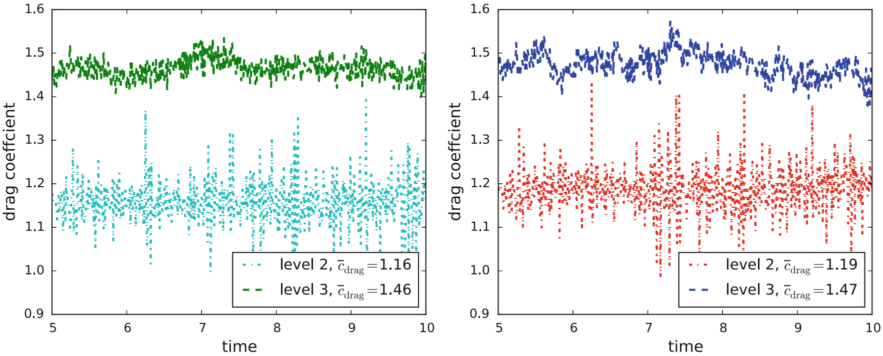

Fig. 8.19 Example 8.282. Computed drag coefficient for the Smagorinsky model with $C_S = 0.005$ (*left*) and VMS0 with $C_S = 0.01$ (*right*)

drag coefficient on the coarser level 2. On the finer grid, level 3, the Smagorinsky model and VMS0 predict mean drag coefficients that are of the same order as the reference value. The models VMS1 and VMSadap underpredict this coefficient still considerably. Figure 8.19 shows that the oscillations of the computed drag coefficients around the mean value are much smaller on the finer grid. □

Chapter 9
Solvers for the Coupled Linear Systems of Equations

Remark 9.1 (Motivation) Many methods for the simulation of incompressible flow problems require the simulation of coupled linear problems for velocity and pressure of the form

$$\mathscr{A}\underline{x} = \begin{pmatrix} A & D \\ B & -C \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{f_p} \end{pmatrix} = \underline{y}, \tag{9.1}$$

with

$$A \in \mathbb{R}^{dN_v \times dN_v},\ D \in \mathbb{R}^{dN_v \times N_p},\ B \in \mathbb{R}^{N_p \times dN_v},\ C \in \mathbb{R}^{N_p \times N_p},$$
$$\underline{u}, \underline{f} \in \mathbb{R}^{dN_v},\ \underline{p}, \underline{f_p} \in \mathbb{R}^{N_p},$$

such that

$$\mathscr{A} \in \mathbb{R}^{(dN_v + N_p)\times(dN_v + N_p)}, \quad \underline{x}, \underline{y} \in \mathbb{R}^{dN_v + N_p}.$$

If $C = 0$, then (9.1) is a linear saddle point problem.

For the Navier–Stokes equations and in particular for time-dependent problems, systems of form (9.1) have to be solved over and over again. The efficient solution of these systems is the core of the overall efficient simulation of incompressible flow problems. □

Remark 9.2 (Contents of This Chapter) The development of fast solvers for problems of type (9.1) is an active field of research. Properties of such systems and classical approaches for their solution are surveyed in Benzi et al. (2005). Solvers for systems of form (9.1) that arise in the discretization of incompressible flow problems are discussed thoroughly in Elman et al. (2014, Chaps. 4 and 9) and Olshanskii and Tyrtyshnikov (2014, Chap. 5.2).

V. John, *Finite Element Methods for Incompressible Flow Problems*, Springer Series in Computational Mathematics 51, DOI 10.1007/978-3-319-45750-5_9

This chapter will concentrate on linear saddle point problems, i.e., $C = 0$ in (9.1), and on the solvers that were used for obtaining the numerical results in this monograph. In addition, some other solvers proposed in the literature are described briefly. □

Remark 9.3 (Sparse Matrices, CSR Storage Format) All matrix blocks are sparse matrices such that $\mathscr{A}$ is sparse, too. Usually, sparse matrices are stored in special formats. The condensed sparse row (CSR) format is very popular. In this format, the non-zero entries are stored row-wise. It is not necessary to order them with respect to the columns. Let $\mathscr{A} \in \mathbb{R}^{m\times n}$ and let nnz be the number of its non-zero entries. Three arrays are needed to store a sparse matrix in the CSR format:

- an array of length nnz where the entries of $\mathscr{A}$ are stored,
- an array of length nnz with integers that contains the column indices of the corresponding entries of the first array,
- an array of length $(m + 1)$ with integers whose i-th entry indicates where the i-th row starts in the other two arrays. The last entry of this array indicates where the $(m + 1)$-st row would start.

For incompressible flow problems, a popular approach is to store each block of A, B, and D individually in this format. □

Remark 9.4 (Some Properties of the System Matrix) In many situations, the system matrix has a more special form compared with (9.1) or some matrix blocks have special properties. In the case of inf-sup stable finite element spaces and the Galerkin finite element method, there are $C = 0$ and $D = B^T$. The second property depends also on the boundary conditions of the concrete problem. For the Stokes equations, the matrix block A is symmetric. Usually, C is also symmetric, e.g., in the PSPG method. □

9.1 Solvers for the Coupled Problems

Remark 9.5 (Sparse Direct Solvers) Easiest to apply are direct solvers for sparse linear systems of equations, so-called sparse direct solvers. Popular solvers are

- MUMPS, Amestoy et al. (2001, 2006),
- pardiso, Schenk et al. (2008),
- umfpack, Davis (2004).

These solvers are particularly efficient for two-dimensional problems and small to medium ($\lesssim$500,000 degrees of freedom) problems. However, they tend to become inefficient for three-dimensional problems, compare Example 7.56. This feature arises from the different sparsity structure of the matrices for discretizations of two- and three-dimensional problems. An internal degree of freedom in three dimensions possesses usually more neighbors than in two dimensions, such that there are more matrix entries and the matrices are less sparse. Sparse direct solvers try to compute a factorization of the matrices with little fill-in, which is easier for matrices where the structure is very sparse.

Direct solvers do not take advantage of situations where a good approximation of the solution is already known. This situation appears in particular in time-dependent problems, discretized with not too large time steps, where the solution from the previous discrete time or some extrapolation of solutions from previous times give usually a good approximation of the solution of the current time.

For ill-conditioned problems, it is known from numerical linear algebra that the numerical results of direct solvers might become inaccurate. Then, it is advised to apply a post-processing with an iterative method for improving the accuracy. □

Remark 9.6 (Iterative Solvers, Krylov Subspace Methods) Let $\underline{r} \in \mathbb{R}^{dN_v+N_p}$, then the space

$$K_k(\underline{r}, \mathscr{A}) := \operatorname{span}\left\{\underline{r}, \mathscr{A}\underline{r}, \ldots, \mathscr{A}^{k-1}\underline{r}\right\}, \quad k \geq 1,$$

is called the Krylov subspace of dimension k that is spanned by $\underline{r}$ and $\mathscr{A}$.

Popular iterative solves for systems of type (9.1) can be found in the family of Krylov subspace methods. Let $\underline{x}^{(0)}$ be some initial approximation of the solution of (9.1), then the residual vector is given by

$$\underline{r}^{(0)} = \underline{y} - \mathscr{A}\underline{x}^{(0)}$$

and the Krylov subspace $K_k\left(\underline{r}^{(0)}, \mathscr{A}\right)$ is considered.

- *GMRES, FGMRES.* The method GMRES (generalized minimal residual), proposed in Saad and Schultz (1986), computes the k-th iterate such that the Euclidean norm of the residual vector is minimized in $\underline{x}^{(0)} + K_k\left(\underline{r}^{(0)}, \mathscr{A}\right)$.
  The advantage of this method is that there is a control on the norm of the residual, since this norm cannot increase. However, one has to store the whole basis of $K_k\left(\underline{r}^{(0)}, \mathscr{A}\right)$. These memory requirements can be afforded usually only for a rather small number k. Also, the cost per iteration (number of floating point operations) increases with each iteration. In practice, one prescribes a maximal number k of iterations, often of the order $10, \ldots, 100$. If the method did not converge after k steps, it is stopped and restarted with the current iterate. This strategy is called GMRES with restart, GMRES(restart).

  As it will be discussed in Sect. 9.2, one has to apply usually a preconditioner for accelerating the speed of convergence. If this preconditioner is not a fixed matrix but a numerical method, then one should use the flexible GMRES methods proposed in Saad (1993). The application of a method can be represented by a matrix, which is usually unknown, but this matrix might change from iteration to iteration. The flexible GMRES method can cope with changes of the preconditioner during the iteration. However, one has to store in the flexible GMRES method the bases of two Krylov spaces, which increases the memory requirements further.

- *CGS.* The main idea of CGS (conjugate gradient squared), proposed in Sonneveld (1989), consists in computing the k-th iterate such that it is orthogonal to $K_k\left(\underline{r}^{(0)}, \mathscr{A}^T\right)$. CGS is a realization of this approach which does not need the knowledge of $\mathscr{A}^T$. This method can be implemented with a so-called short recurrence, i.e., one needs only a few arrays and the needed number of arrays does not increase during the iteration. Hence, the memory requirements are much less compared with GMRES. However, the Euclidean norm of the residual vector is not minimized and it can happen that this norm increases hugely during the iteration. In this sense, an irregular convergence behavior is possible.
- *BiCGStab.* The BiCGStab (bi-conjugate gradient stabilized) method, developed in van der Vorst (1992), relies on the same principle as the CGS method. It is a modification of the CGS method that is usually more stable.

To accelerate the convergence of iterative solvers, one has to apply so-called preconditioners, see Sect. 9.2.

Available libraries providing iterative solvers together with preconditioners for linear saddle point problems are

- PETSc, Balay et al. (2016),
- Trilinos, Trilinos (2016).

□

Remark 9.7 (Strategies for Solving (9.1) *Iteratively)* For the Navier–Stokes equations, systems of form (9.1) arise in the Picard or Newton iteration, see Sect. 6.3. From experience, it is known that it is often inefficient to solve (9.1) to high accuracy if an iterative solver is used. For this reason, one performs inexact solutions. This strategy means that one prescribes as stopping criteria a small number of maximal iterations or a small reduction factor for the Euclidean norm of the residual vector, e.g., 10 to gain one digit. After having satisfied one of the stopping criteria, one proceeds with the next Picard or Newton step. The performance of this strategy is illustrated in Examples 6.47 and 7.56. □

9.2 Preconditioners for Iterative Solvers

Remark 9.8 (Preconditioners) It is well known that the performance of iterative solvers can be improved considerably if so-called preconditioners are used. Preconditioners are approximations of $\mathscr{A}^{-1}$ that can be computed comparatively efficiently. These approximations might be a fixed matrix or a numerical (iterative) method or a combination of both.

One distinguishes between left and right preconditioners. Denote the preconditioner by $\mathcal{M}^{-1}$. Then, for left preconditioning, one considers instead of (9.1) the system

$$\mathcal{M}^{-1}\mathscr{A}\underline{x} = \mathcal{M}^{-1}\underline{y}$$

in the iterative solver. Iterative solvers with right preconditioner solve the problem

$$\mathscr{A}\mathcal{M}^{-1}\underline{z} = \underline{y}, \quad \underline{x} = M^{-1}\underline{z}.$$

Iterative solvers with left preconditioner might behave differently than the same solver with the same preconditioner used as right preconditioner. For instance, left and right preconditioned GMRES minimize different residuals, e.g., see Saad (2003, Sect. 9.3.4) for details. □

9.2.1 Incomplete Factorizations

Remark 9.9 (On Incomplete Factorizations) Applying a LU factorization to a sparse matrix A, the factors L and U are generally considerably denser than A. In an incomplete LU factorization (ILU), one stores factors L and U with the same sparsity pattern as A, i.e., L is stored with a sparsity pattern that corresponds to the strict lower triangle of A and U with a sparsity pattern that corresponds to the pattern of the upper triangle of A. Performing the algorithm of the standard LU decomposition (without pivoting), one neglects all entries that do not fit in the prescribed sparsity pattern. In this way, one obtains an incomplete decomposition

$$A = LU + E \tag{9.2}$$

with the error matrix E.

It is noted, e.g., in Dahl and Wille (1992), that avoiding pivoting in the application of the ILU factorization of $\mathcal{A}$ does not lead necessarily to a failure of this method. First, one should order the degrees of freedom in the way that the pressure degrees of freedom come last, like in (9.1). Second, one has to provide the sparse matrix C, where the sparsity pattern contains all connections of pressure degrees of freedom, like for an inf-sup stabilized method, and all entries of C are initialized with zero. Under these two conditions, it was observed in Dahl and Wille (1992) that a division by zero does not occur because the zero entries at the main diagonal vanish during the incomplete factorization.

A so-called ILUT(l, τ) method, Saad (2003, Sect. 10.4.1), allows to accept entries computed in the LU factorization whose absolute value is larger than $\tau \geq 0$, even if they do not fit in the sparsity pattern. Only the l largest (with respect to the absolute value) entries in the L part of the considered row and the l largest entries in the U part of this row, plus the diagonal entry, are kept. In this way, an incomplete decomposition of form (9.2) is obtained, with presumably smaller entries of the error matrix and without the full fill-in of a complete LU decomposition. □

Remark 9.10 (An Incomplete LU Preconditioner with Fill-in) In Konshin et al. (2015), an ILU preconditioner for the solution of linear saddle point problems arising in the discretization of the time-dependent Navier–Stokes equations was proposed. This preconditioner considers the factorization of the global matrix $\mathscr{A}$, which does not use pivoting, and which is based on a factorization of the form

$$\mathscr{A} = LU + LR_u + R_l U + E,$$

where R_u is a strictly upper triangular matrix, R_l is a strictly lower triangular matrix, and E is again the error matrix. Given two parameters $\tau_1 \geq \tau_2 > 0$, the decomposition is constructed in such a way that the off-diagonal elements of U and L are either zero or they have an absolute value that is greater than τ_1, while the off-diagonals in R_u, R_l are either zero or their absolute value belongs to $(\tau_2, \tau_1]$. Consequently, the entries of the error matrix E are of order $\mathcal{O}(\tau_2)$. As explained in Konshin et al. (2015), the fill-in of this method is determined by τ_1 and the quality of the resulting preconditioner is ruled by τ_2 once $\tau_2 = \mathcal{O}(\tau_1)$.

For the proposed preconditioner to work efficiently, in a first step, an appropriate scaling of the matrix $\mathscr{A}$ is necessary. It was shown in Konshin et al. (2015) that this preconditioner, with suitable values τ_1 and τ_2 worked efficiently for laminar, three-dimensional, time-dependent Navier–Stokes problems for a wide range of the viscosity. □

9.2.2 A Coupled Multigrid Method

Remark 9.11 (Multigrid Methods) Multigrid methods use a hierarchy of grids or levels for the solution of linear systems of equations. On each level, an appropriate equation is solved approximately. Only on the coarsest level, which is called level 0, the equation might be solved with high accuracy. Performing a multigrid method requires the transport of information between subsequent levels. The transport from finer to coarser levels is called restriction and the other way is called prolongation. There are several introductions to multigrid methods, e.g., Hackbusch (1985), Trottenberg et al. (2001) or Briggs et al. (2000), which describe their main ideas in detail.

Here, the components of a multigrid method for problem (9.1) will be described. This method solves (9.1) for both types of unknowns, $\underline{u}$ and $\underline{p}$ together and therefore it is called coupled multigrid method. It is defined by

- the grid hierarchy, Remark 9.12,
- the definition of the system matrix on coarser levels, Remark 9.13,
- the grid transfer operators
  - function prolongation, Remark 9.14,
  - defect restriction, Remark 9.18,
  - function restriction, Remark 9.19,
- the smoother on finer levels, Remark 9.20,
- the coarse grid solver, Remark 9.24. □

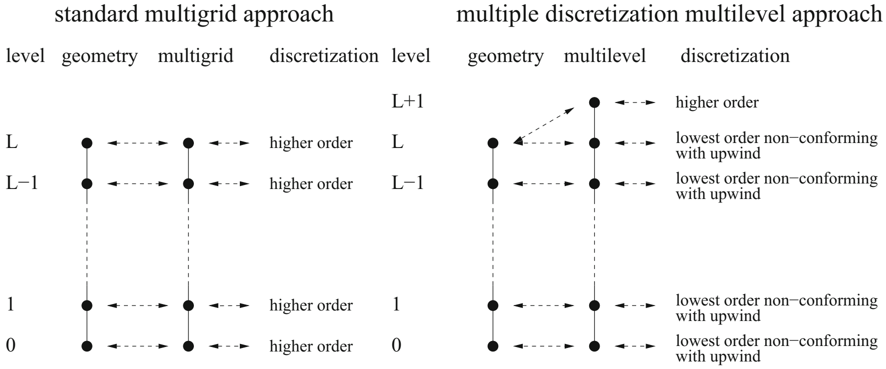

Fig. 9.1 The standard multigrid approach (*left*) and the multiple discretization multilevel approach for higher order discretizations (*right*)

Remark 9.12 (The Grid Hierarchy) There are several possibilities for defining a multigrid hierarchy. The standard way consists in using the same levels for the multigrid hierarchy as they are given by the hierarchy of the geometrical grids and to use the same discretization on all levels. However, the multigrid hierarchy might have more levels than the geometric grid hierarchy and other discretizations than on the finest level might be used on coarser levels. A representative of this class is the multiple discretization multilevel approach shown in Fig. 9.1.

In numerical studies of Example D.5 in John and Matthies (2001) and of Example D.6 in John (2002), difficulties were observed with the standard multigrid approach for solving linear saddle point problems arising in some higher order finite element discretizations. Higher order finite element discretizations give quite accurate results for the quantities of interest, compare Example 6.36. In contrast, lowest order non-conforming discretization with upwind stabilization, see Remark 5.58, were rather inaccurate but the standard multigrid approach has been proved as a very efficient solver, see also John and Tobiska (2000). This situation led to the idea of constructing a multilevel method for higher order finite element discretizations that is based on a multilevel method for stable lowest order non-conforming finite element discretizations. This is just the multiple discretization multilevel method from Fig. 9.1. In this method, the multilevel hierarchy possesses one level more than the geometric grid hierarchy. On the finest geometric grid, level L, two discretizations are applied. One of them, which forms the finest level of the multilevel hierarchy, is the discretization of interest, i.e., the higher order discretization. The second discretization on the geometric level L is a lowest order non-conforming discretization with upwind stabilization. On all coarser geometric levels, also a stabilized lowest order non-conforming discretization is applied.

There are several possibilities for defining the sequence of grids, the so-called cycles, in a multigrid method. The most common multigrid cycles are depicted in Fig. 9.2. From the point of view of the number of floating point operations, the cheapest cycle is the V-cycle and the W-cycle is most expensive. With respect

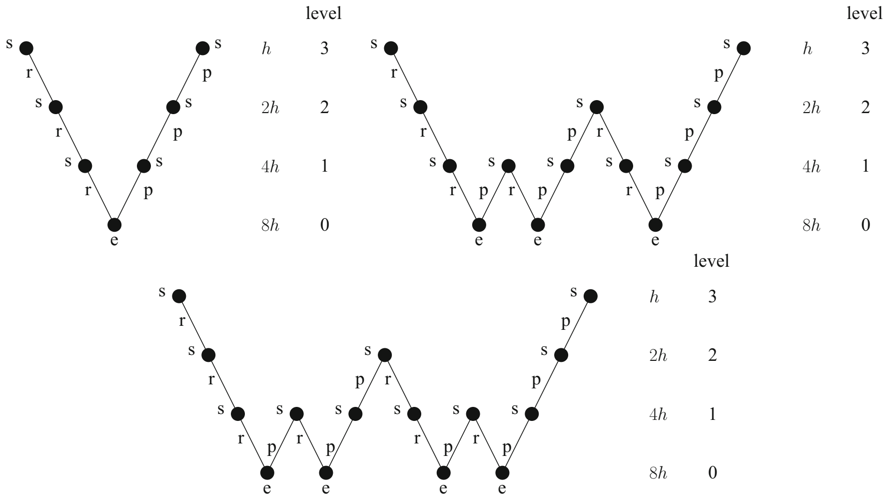

Fig. 9.2 Multigrid V-cycle, F-cycle, W-cycle (*top left, top right, bottom*), s—smoothing, r—restriction, p—prolongation, e—exact solver

to stability, the situation is vice versa. Often, the F-cycle is a good compromise between computational costs and stability. □

Remark 9.13 (The Definition of the Linear Systems on Coarser Levels) The so-called Galerkin projection defines the system matrix on level $(l-1)$ with the help of the system matrix on level l and the prolongation and restriction operators. However, in this approach, it is not possible to apply different discretizations on different levels, which is the idea of the multiple discretization multilevel method. In addition, it seems to be easier from the point of view of implementation to directly assemble the system matrices on coarser levels in the same way as the system matrix on the finest level. Thus, the direct assembling is often used in practice. □

Remark 9.14 (The Function Prolongation) Performing a multigrid method requires the transfer of finite element functions from one level of the multigrid hierarchy to the next finer level. It might happen, e.g., in the case of the multiple discretization multilevel method, that the finite element spaces on subsequent levels are not nested. The transfer operator has to cope with this situation. In Schieweck (2000), a transfer operator was introduced and analyzed that allows the transfer between almost arbitrary finite element spaces. This operator will be described in some detail, for simplicity for scalar finite element functions. It is based on the concept of nodal functionals, see Remark B.18, and the local basis, see Remark B.20.

Consider the transfer (prolongation) from a finite element space V_{l-1}^h to a finite element space V_l^h. Let $\mathcal{T}_{l-1}$ and $\mathcal{T}_l$ be the corresponding triangulations of the domain Ω such that $\mathcal{T}_l$ originates either from a refinement of $\mathcal{T}_{l-1}$ or $\mathcal{T}_{l-1} = \mathcal{T}_l$. The second case is relevant in the multiple discretization multilevel method for $l = L + 1$, see Fig. 9.1.

Let S_l^h be a discontinuous finite element space defined on $\mathcal{T}_l$

$$S_l^h = \left\{ w \in L^2(\Omega) \, : \, w|_K \in S_l^h(K) \; \forall \; K \in \mathcal{T}_l \right\}.$$

The choice of the local spaces $S_l^h(K)$ depends on V_{l-1}^h and V_l^h. It has to be done in such a way that the inclusion

$$V_{l-1}^h + V_l^h \subset S_l^h \tag{9.3}$$

holds. From the practical point of view, the spaces $S_l^h(K)$ are not needed for implementing the transfer operator. From the theoretical point of view, it can be proved that appropriate spaces $S_l^h(K)$ always exist for triangulations consisting of simplices, see John et al. (2002).

The transfer operator for the prolongation is defined with the help of the global nodal functionals, see Remark B.23,

$$P_{l-1}^l : S_l^h \to V_l^h \quad P_{l-1}^l\left(w^h\right) = \sum_{i=1}^{\dim\left(V_l^h\right)} N_{l,i}\left(w^h\right) \varphi_{l,i}^h, \tag{9.4}$$

where $\left\{\varphi_{l,i}^h\right\}_{i=1}^{\dim\left(V_l^h\right)}$ is a finite element basis of V_l^h and $N_{l,i}\left(w^h\right)$ is the global nodal functional at level l such that $N_{l,i}\left(\varphi_{l,j}^h\right) = \delta_{ij}$. From the inclusion (9.3), it follows that the operator (9.4) is defined especially for functions from V_{l-1}^h. □

Remark 9.15 (Evaluation of the Global Functionals with Local Functionals) Let $\left\{\varphi_{l-1,i}^h\right\}_{i=1}^{\dim\left(V_{l-1}^h\right)}$ be a finite element basis of V_{l-1}^h and let

$$w_{l-1}^h = \sum_{i=1}^{\dim\left(V_{l-1}^h\right)} w_{l-1,i}\varphi_{l-1,i}^h \in V_{l-1}^h.$$

For evaluating the coefficient of $\varphi_{l,i}^h$ for the prolongation of the function w_{l-1}^h, one has to compute

$$\begin{aligned} N_{l,i}\left(w_{l-1}^h\right) &= \frac{1}{\operatorname{card}\left(\mathcal{T}_{l,i}\right)} \sum_{K \in \mathcal{T}_{l,i}} N_{l,i}^K\left(w_{l-1}^h|_K\right) \\ &= \frac{1}{\operatorname{card}\left(\mathcal{T}_{l,i}\right)} \sum_{K \in \mathcal{T}_{l,i}} \sum_{j=1}^{\dim\left(V_{l-1}^h\right)} w_{l-1,j} N_{l,i}^K\left(\varphi_{l-1,j}^h|_K\right). \end{aligned}$$

□

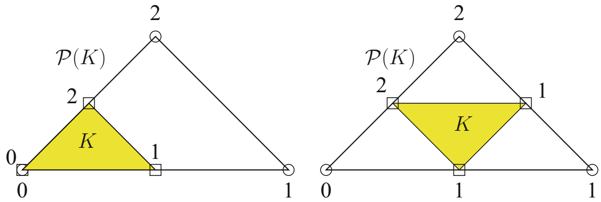

Fig. 9.3 Red refined triangles, $P_1(\mathcal{P}(K)) \to P_1(K)$

Example 9.16 (Evaluation of Local Functionals) Some concrete situations for the evaluation of $N_{l,i}^K\left(\varphi_{l-1,j}^h|_K\right)$ are described in detail. For simplicity, the presentation is restricted to two-dimensional finite elements. Let $\mathcal{P}(K) \in \mathcal{T}_{l-1}$ be the parent mesh cell of $K \in \mathcal{T}_l$. The local degrees of freedom of $\mathcal{P}(K)$ are represented by circles in Figs. 9.3, 9.4, 9.5, and 9.6 and the local degrees of freedom of K by squares.

- *Red refined triangles*, $P_1(\mathcal{P}(K)) \to P_1(K)$. Consider the two situations given in Fig. 9.3. For this finite element, the local nodal functionals are point values, i.e., $N_{l,i}^K\left(\varphi_{l-1,j}^h|_K\right)$ is the value of $\varphi_{l-1,j}^h$ at the position of the local degree of freedom i in K. One obtains the following values

	Figure 9.3, left			Figure 9.3, right		
	$\varphi_{l-1,0}^h\|_K$	$\varphi_{l-1,1}^h\|_K$	$\varphi_{l-1,2}^h\|_K$	$\varphi_{l-1,0}^h\|_K$	$\varphi_{l-1,1}^h\|_K$	$\varphi_{l-1,2}^h\|_K$
$N_{l,0}^K$	1	0	0	0.5	0.5	0
$N_{l,1}^K$	0.5	0.5	0	0	0.5	0.5
$N_{l,2}^K$	0.5	0	0.5	0.5	0	0.5

  It turns out for the prolongation that this is just the standard inclusion.
- *Red refined triangles*, $P_1^{\mathrm{nc}}(\mathcal{P}(K)) \to P_1^{\mathrm{nc}}(K)$. Consider again two situations, see Fig. 9.4. In this case, the local nodal functionals are given by integral mean values at the edges of K, e.g.,

$$N_{l,0}^K\left(\varphi_{l-1,j}^h|_K\right) = \frac{1}{\|\boldsymbol{x}_0 - \boldsymbol{x}_1\|_2} \int_{\boldsymbol{x}_0}^{\boldsymbol{x}_1} \varphi_{l-1,j}^h|_K \, ds.$$

  One obtains for the local nodal functionals

	Figure 9.4, left			Figure 9.4, right		
	$\varphi_{l-1,0}^h\|_K$	$\varphi_{l-1,1}^h\|_K$	$\varphi_{l-1,2}^h\|_K$	$\varphi_{l-1,0}^h\|_K$	$\varphi_{l-1,1}^h\|_K$	$\varphi_{l-1,2}^h\|_K$
$N_{l,0}^K$	1	−0.5	0.5	0.5	0.5	0
$N_{l,1}^K$	0.5	0	0.5	0	0.5	0.5
$N_{l,2}^K$	0.5	−0.5	1	0.5	0	0.5

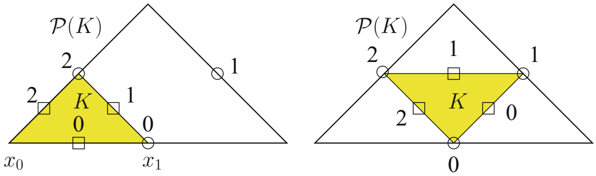

Fig. 9.4 Red refined triangles, $P_1^{\text{nc}}(\mathcal{P}(K)) \to P_1^{\text{nc}}(K)$

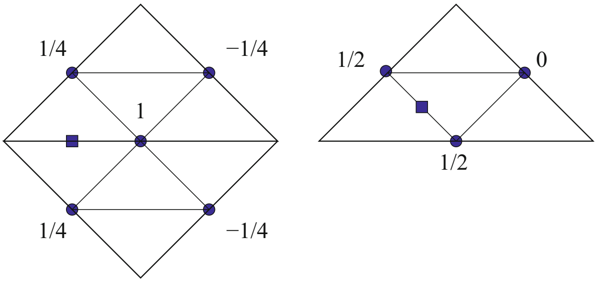

Fig. 9.5 Red refined triangles, the weights of the prolongation for $P_1^{\text{nc}}(\mathcal{P}(K)) \to P_1^{\text{nc}}(K)$

Applying these local nodal functionals in the prolongation operator (9.4), one gets a standard averaging operator, see Fig. 9.5. In this picture, the square denotes the degree of freedom of V_l^h whose value has to be computed and the balls stand for the nodes of V_{l-1}^h. The numbers give the weights that have to be applied to the coefficients of the function from V_{l-1}^h corresponding to these nodes. It can be seen that the prolongated value in the left picture of Fig. 9.5 is just the average in this point of the values of the finite element function of V_{l-1}^h restricted to the triangles in $\mathcal{T}_{l-1}$.

- *No refined affinely mapped quadrilaterals*, $Q_1^{\text{rot}}(\mathcal{P}(K)) \to Q_2(K)$. In this case, it is $\mathcal{P}(K) = K$. One starts by changing the basis in $Q_1^{\text{rot}}\left(\hat{K}\right)$. It is straightforward to check that for the local basis of $Q_1^{\text{rot}}\left(\hat{K}\right)$ given in (B.22) it holds

$$\frac{1}{\left|\hat{E}_i\right|}\int_{\hat{E}_i} \hat{\varphi}_{l-1,j}^h \, ds = \delta_{ij},$$

where the edges $\hat{E}_i$ of $\hat{K}$ are numbered counter-clockwise, starting with the bottom edge. Let K be an arbitrary mesh cell with an affine reference transformation and let $(\hat{x}_1, \hat{x}_2) \in \hat{K}$ be transformed to $(x_1, x_2) \in K$. Then, it holds for the transformation of the basis functions that $\hat{\varphi}^h(\hat{x}_1, \hat{x}_2) = \varphi^h(x_1, x_2)$. Thus, the values of the transformed basis functions can be easily computed by values of the reference basis functions in $\hat{K}$. The affine reference transformation leads to a situation as presented in Fig. 9.6. The local nodal functionals of $Q_2(K)$ are

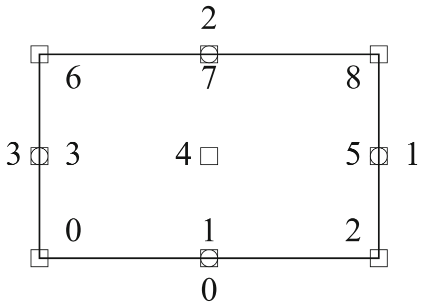

Fig. 9.6 No refined affine mapped quadrilaterals, $Q_1^{\text{rot}}(\mathcal{P}(K)) \to Q_2(K)$

defined as point values of the local basis functions of $Q_1^{\text{rot}}(K)$. As pointed out, the evaluation of these point values can be done in $\hat{K}$ which gives, independent of K,

	$\varphi_{l-1,0}^h\|_K$	$\varphi_{l-1,1}^h\|_K$	$\varphi_{l-1,2}^h\|_K$	$\varphi_{l-1,3}^h\|_K$
$N_{l,0}^K$	$3/4$	$-1/4$	$-1/4$	$3/4$
$N_{l,1}^K$	$9/8$	$-1/8$	$1/8$	$-1/8$
$N_{l,2}^K$	$3/4$	$3/4$	$-1/4$	$-1/4$
$N_{l,3}^K$	$-1/8$	$1/8$	$-1/8$	$9/8$
$N_{l,4}^K$	$1/4$	$1/4$	$1/4$	$1/4$
$N_{l,5}^K$	$-1/8$	$9/8$	$-1/8$	$1/8$
$N_{l,6}^K$	$-1/4$	$-1/4$	$3/4$	$3/4$
$N_{l,7}^K$	$1/8$	$-1/8$	$9/8$	$-1/8$
$N_{l,8}^K$	$-1/4$	$3/4$	$3/4$	$-1/4$

In practice, the same values are used also in the situation that the reference map F_K is bilinear.

These examples demonstrate that the values of the local nodal functionals $N_{l,i}^K\left(\varphi_{l-1,j}^h|_K\right)$ are, in general, the same for a large number of mesh cells. The values can be computed in a pre-processing step and stored in a data base. In computing the prolongation, only local matrix-vector products have to be performed with these values. This strategy is used to accelerate the computation of the prolongation (9.4).
□

Remark 9.17 (Update After the Prolongation) Let $\left(\boldsymbol{u}_l^h, p_l^h\right)$ be the current iterate at the multigrid level l and $\left(\delta\boldsymbol{u}_l^h, \delta p_l^h\right)$ be the update computed on level $l-1$. Then, the next iterate on level l is given by $\left(\boldsymbol{u}_l^h, p_l^h\right) + \kappa_l P_{l-1}^l\left(\delta\boldsymbol{u}_l^h, \delta p_l^h\right)$, where κ_l is a damping factor. In practice, velocity and pressure have to be prolongated separately using different operators. Often, the choice $\kappa_l = 1$ works well, but sometimes smaller values of κ_l are beneficial for the efficiency and robustness of the method. □

Remark 9.18 (The Defect Restriction) The definition of the operator for the defect restriction $(R')_l^{l-1} : \left(V_l^h\right)' \to \left(V_{l-1}^h\right)'$ uses the prolongation operator given in (9.4). Let $d_l \in \left(V_l^h\right)'$ be a given defect functional, then its restriction to $\left(V_{l-1}^h\right)'$ is defined by

$$\int_\Omega \left(R'\right)_l^{l-1} (d_l)\, \varphi_{l-1}^h \, d\boldsymbol{x} = \int_\Omega d_l P_{l-1}^l \left(\varphi_{l-1}^h\right) \, d\boldsymbol{x} \quad \forall \, \varphi_{l-1}^h \in V_{l-1}^h.$$

Since the prolongation operator turns out to be a standard operator in many situations, see Example 9.16, the same holds for the defect restriction operator. □

Remark 9.19 (The Function Restriction) As described in Remark 9.13, the system matrix on coarser levels will be obtained by direct assembling. The assembling of the convective term of the linearized Navier–Stokes equations has as parameter the current finite element velocity $\boldsymbol{u}_{\text{old}}^h$. An appropriate restriction of this function has to be available on all levels. A restriction operator $R_l^{l-1} : V_l^h \to V_{l-1}^h$ that maps a finite element function from the finite element space connected to level l in the multilevel hierarchy to a finite element function connected to level $(l-1)$ is necessary. A possible definition of this operator is based on local projections in the sense of $L^2(K)$ and averaging.

Denote the bases of V_l^h and V_{l-1}^h by $\{\varphi_{l,i}^h\}_{i=1}^{\dim\left(V_l^h\right)}$ and $\{\varphi_{l-1,i}^h\}_{i=1}^{\dim\left(V_{l-1}^h\right)}$, respectively, and let $\boldsymbol{v}_l^h \in V_l^h$ be given. The goal consists in computing a function $R_l^{l-1}\left(\boldsymbol{v}_l^h\right)$. Using the bases, $\boldsymbol{v}_l^h$ and its restriction can be represented by

$$\boldsymbol{v}_l^h = \sum_{i=1}^{\dim\left(V_l^h\right)} v_{l,i}\varphi_{l,i}^h \quad \text{and} \quad R_l^{l-1}\left(\boldsymbol{v}_l^h\right) = \sum_{i=1}^{\dim\left(V_{l-1}^h\right)} v_{l-1,i}\varphi_{l-1,i}^h.$$

Consider a mesh cell K on the geometric grid that is connected with V_{l-1}^h and assume that K possesses an affine reference transformation. Local values of the unknown coefficients $v_{l-1,i}$ are determined by the local $L^2(K)$ projection

$$\sum_{i=1}^{\text{card}\left(\mathcal{I}_l\left(K,V_l^h\right)\right)} v_{l,i}|_K \left(\varphi_{l,i}^h, \varphi_{l-1,j}^h\right)_K = \sum_{i=1}^{\text{card}\left(\mathcal{I}_{l-1}\left(K,V_{l-1}^h\right)\right)} v_{l-1,i}|_K \left(\varphi_{l-1,i}^h, \varphi_{l-1,j}^h\right)_K$$

for all $j \in \mathcal{I}_{l-1}\left(K, V_{l-1}^h\right)$, where $\mathcal{I}_l\left(V_l^h\right)$ is the set of degrees of freedom of V_l^h and

$$\mathcal{I}_l\left(K, V_l^h\right) = \left\{ i \in \mathcal{I}_l\left(V_l^h\right) \, : \, \text{supp}\left(\varphi_{l,i}^h\right) \cap \mathring{K} \neq \emptyset \right\}$$

is the set of local degrees of freedom with respect to the mesh cell K. The transformation (B.32) to the reference cell $\hat{K}$ yields

$$\begin{aligned}
&\sum_{i=1}^{\operatorname{card}\left(\mathcal{I}_l\left(K,V_l^h\right)\right)} v_{l,i}|_K \int_{\hat{K}} \hat{\varphi}_{l,i}^h \hat{\varphi}_{l-1,j}^h \left|\det J_K\left(\hat{\boldsymbol{x}}\right)\right| \, d\hat{\boldsymbol{x}} \\
&= \sum_{i=1}^{\operatorname{card}\left(\mathcal{I}_{l-1}\left(K,V_{l-1}^h\right)\right)} v_{l-1,i}|_K \int_{\hat{K}} \hat{\varphi}_{l-1,i}^h \hat{\varphi}_{l-1,j}^h \left|\det J_K\left(\hat{\boldsymbol{x}}\right)\right| \, d\hat{\boldsymbol{x}}
\end{aligned}$$

for all $j \in \mathcal{I}_{l-1}\left(K, V_{l-1}^h\right)$. Since $\left|\det J_K\left(\hat{\boldsymbol{x}}\right)\right|$ is constant, this relation simplifies to

$$\begin{aligned}
&\sum_{i=1}^{\operatorname{card}\left(\mathcal{I}_l\left(K,V_l^h\right)\right)} v_{l,i}|_K \int_{\hat{K}} \hat{\varphi}_{l,i}^h \hat{\varphi}_{l-1,j}^h \, d\hat{\boldsymbol{x}} \\
&= \sum_{i=1}^{\operatorname{card}\left(\mathcal{I}_{l-1}\left(K,V_{l-1}^h\right)\right)} v_{l-1,i}|_K \int_{\hat{K}} \hat{\varphi}_{l-1,i}^h \hat{\varphi}_{l-1,j}^h \, d\hat{\boldsymbol{x}}
\end{aligned} \tag{9.5}$$

for all $j \in \mathcal{I}_{l-1}\left(K, V_{l-1}^h\right)$. Hence, one obtains a linear system of equations of the form

$$G v_l|_K = M v_{l-1}|_K.$$

Thus, the local values of the unknown coefficients are given by

$$v_{l-1}|_K = M^{-1} G v_l|_K = R v_l|_K. \tag{9.6}$$

The matrix R is independent of K. Hence, for all other mesh cells whose basis on the reference mesh cell has the same form as for K, it can be also used. This situation occurs very often. For instance, if the grids are uniformly refined and the same finite element space is used on every level, the matrix R is needed for each mesh cell on each level. This matrix R can be computed once and then stored in a data base. Then, only a local matrix-vector product has to be computed in (9.6), which leads to an efficient algorithm. The final restriction is computed by an averaging

$$v_{l-1,i} = \frac{1}{\operatorname{card}\left(\mathcal{T}_{l-1,i}\right)} \sum_{K \in \mathcal{T}_{l-1,i}} v_{l-1,i}|_K.$$

For mesh cells with a non-affine reference transformation, one can use for simplicity also (9.5) such that they are handled in the same way as mesh cells with an affine transformation. This approach can be considered as a function restriction that is a local $L^2(\hat{K})$ projection on the reference mesh cell and that is an approximation of a $L^2(K)$ projection on the original mesh cell. □

Remark 9.20 (Smoothing by Solving Local Problems, the Vanka Smoother) Coupled multigrid methods for the linearized Navier–Stokes equations are usually used with local smoothers, so-called Vanka-type smoothers, Vanka (1986). Vanka-type smoothers can be considered as block Gauss–Seidel methods. Let $\mathcal{V}^h$ and $\mathcal{Q}^h$ be the set of velocity and pressure degrees of freedom, respectively. These sets are decomposed into

$$\mathcal{V}^h = \cup_{j=1}^J \mathcal{V}_j^h, \quad \mathcal{Q}^h = \cup_{j=1}^J \mathcal{Q}_j^h. \tag{9.7}$$

The subsets are not required to be disjoint.

Let $\mathscr{A}_j$ be the block of the matrix $\mathscr{A}$ that is connected with the degrees of freedom of $\mathcal{W}_j^h = \mathcal{V}_j^h \cup \mathcal{Q}_j^h$, i.e., the intersection of the rows and columns of $\mathscr{A}$ with the global indices belonging to $\mathcal{W}_j^h$,

$$\mathscr{A}_j = \begin{pmatrix} A_j & D_j \\ B_j & 0 \end{pmatrix} \in \mathbb{R}^{\dim\left(\mathcal{W}_j^h\right)\times\dim\left(\mathcal{W}_j^h\right)}.$$

Similarly, denote by $(\cdot)_j$ the restriction of a vector to the rows corresponding to the degrees of freedom in $\mathcal{W}_j^h$. Each smoothing step with a Vanka-type smoother consists in a loop over all sets $\mathcal{W}_j^h$, where for each $\mathcal{W}_j^h$ a local system of equations connected with the degrees of freedom in this set is solved. The local solutions are updated in a Gauss–Seidel manner. The Vanka smoother computes new velocity and pressure values by

$$\begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix}_j := \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix}_j + \mathscr{A}_j^{-1} \left(\begin{pmatrix} \underline{f} \\ \underline{f_p} \end{pmatrix} - \mathscr{A} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} \right)_j.$$

The local systems of equations are usually solved with a direct solver.

A general strategy for choosing the sets $\mathcal{V}_j^h$ and $\mathcal{Q}_j^h$ is as follows. First, one picks some pressure degrees of freedom that define $\mathcal{Q}_j^h$. Second, $\mathcal{V}_j^h$ is formed by all velocity degrees of freedom that are connected with the pressure degrees of freedom from $\mathcal{Q}_j^h$ by entries in the sparsity pattern of the matrix B. □

Example 9.21 (The Mesh-Cell-Oriented Vanka Smoother for Pairs with Discontinuous Finite Element Pressure) For the mesh-cell-oriented Vanka smoother, $\mathcal{Q}_j^h$ is defined by all pressure degrees of freedom which are connected to the mesh cell j. For this type of Vanka smoother, J coincides with the number of mesh cells. The mesh-cell-oriented Vanka smoother is applied usually for discretizations with discontinuous pressure approximation, i.e., $Q^h \in \left\{P_0, Q_0, P_1^{\mathrm{disc}}, P_2^{\mathrm{disc}}\right\}$. Then, $\mathcal{V}_j^h$ consists of all velocity degrees of freedom that are connected to the mesh cell j.

Table 9.1 Degrees of freedom for the local systems of the mesh-cell-oriented Vanka smoother (velocity: each component)

	2d			3d		
	Velocity	Pressure	Total	Velocity	Pressure	Total
Q_1^{rot}/Q_0	4	1	9	6	1	19
Q_2/P_1^{disc}	9	3	21	27	4	85
Q_3/P_2^{disc}	16	6	38	64	10	202
P_1^{nc}/P_0	3	1	7	4	1	13

Using a discontinuous discrete pressure, the local matrix of the mesh-cell-oriented Vanka smoother can be generated on the current mesh cell. The size of the local systems is known a priori and it is given for several discretizations in Table 9.1. □

Example 9.22 (The Pressure-Node-Oriented Vanka Smoother for Pairs with Continuous Finite Element Pressure) For spaces with continuous pressure, usually a decomposition is applied where Q_j^h is defined by a single pressure degree of freedom, i.e., $\dim Q_j^h = 1$. For this so-called pressure-node-oriented Vanka smoother, the number of subsets J in decomposition (9.7) is equal to the number of pressure degrees of freedom.

For a continuous pressure approximation, a pressure degree of freedom on a given mesh cell K is in general connected to velocity degrees of freedom on other mesh cells.

The size of the local systems for the pressure-node-oriented Vanka smoother applied in discretizations with continuous pressure approximation depends on the particular pressure degree of freedom and on the given grid. In addition, the size of the local systems cannot be bounded a priori if adaptive grid refinement is used since it depends on the maximal number of neighbor cells of K. A neighbor is a mesh cell K_1 with $K \cap K_1 \neq \emptyset$ and the maximal number of neighbor cells of a mesh cell K can increase on adaptively refined grids. To illustrate the size of the local systems, concrete values are given in Table 9.2 for the typical situations depicted in Fig. 9.7. It can be observed that these sizes are considerably larger than for the mesh-cell-oriented Vanka smoother for discretizations with discontinuous pressure, compare Table 9.1. In addition, the number of local systems that must be solved in each smoothing step is in general larger compared with the mesh-cell-oriented Vanka smoother since there are generally more pressure degrees of freedom than mesh cells. □

Remark 9.23 (Damping of the Smoother Iterate) Sometimes it is beneficial to damp the smoother iterate. Let $\left(\underline{u}_l, \underline{p}_l\right)$ be the current iterate at the multigrid level l and $\left(\delta\underline{u}_l, \delta\underline{p}_l\right)$ be the update computed by one iteration of the smoother. Then, the new

Table 9.2 Degrees of freedom for the local systems of the pressure-node-oriented Vanka smoother for the pressure degrees of freedom from Fig. 9.7 (velocity: each component)

	2d			3d		
	Velocity	Pressure	Total	Velocity	Pressure	Total
Q_2/Q_1	25	1	51	125	1	376
Q_3/Q_2	49	1	99	343	1	1030
P_2/P_1	19	1	39	65	1	196
P_3/P_2	37	1	75	175	1	526

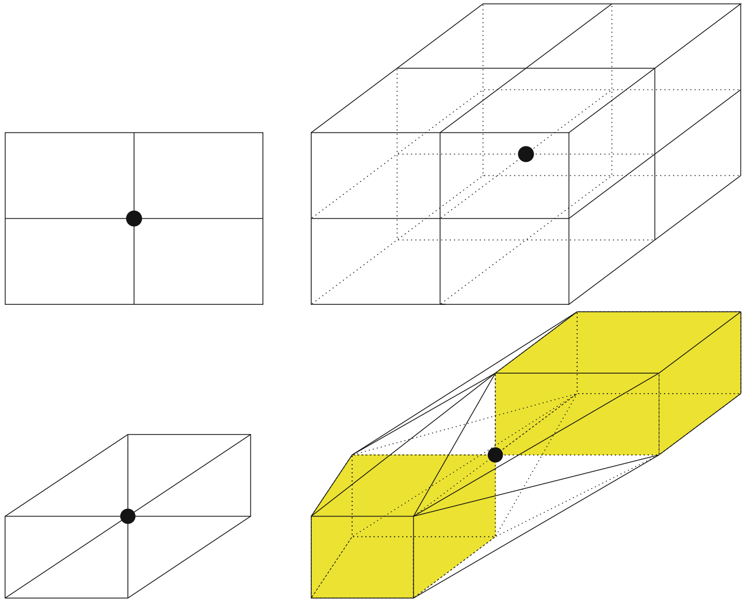

Fig. 9.7 Degree of freedom for which the size of the local systems in the pressure-node-oriented Vanka smoother is given in Table 9.2 (*bottom right*: 6 tetrahedra in two directions (*coloured*) and 2 tetrahedra in six directions, i.e., 24 tetrahedra are connected with this pressure degree of freedom)

iterate is computed by $\left(\underline{u}_l, \underline{p}_l\right) + \omega_l \left(\delta \underline{u}_l, \delta \underline{p}_l\right)$. The damping parameter can be chosen differently on all levels of the multigrid hierarchy. Instead of choosing a fixed damping factor ω_l, an automatic step length control as proposed in John and Tobiska (2000) is sometimes helpful. □

Remark 9.24 (The Coarse Grid Solver) As coarse grid solver, one can apply the Vanka smoother or a direct solver. If the Vanka smoother is used, then usually more iterations are performed than on finer grids, often of the order 10–50. □

Remark 9.25 (Coupled Multigrid Methods as Solver and as Preconditioner) Coupled multigrid methods can be used as solver or as preconditioner. The experience is that the application as preconditioner is more efficient, e.g., see John (2002). Since the preconditioner is in this case not a fixed matrix but a method, it should be used in a flexible Krylov subspace method, e.g., in FGMRES, see Remark 9.6. The right preconditioned FGMRES was used, e.g., in John (2002, 2006) and in the simulations presented in this monograph. □

9.2.3 Preconditioners Treating Velocity and Pressure in a Decoupled Way

Remark 9.26 (Motivation) The straightforward application of many standard schemes for the solution or preconditioning of linear systems of equations becomes difficult because of the block structure of the system matrix $\mathscr{A}$ from (9.1). For this reason, preconditioners have been developed where individual linear systems of equations for the pressure and for each component of the velocity have to be solved. These individual systems do not possess a block structure, which enables the application of standard solvers and preconditioners. This class of methods is sometimes called segregated schemes. □

Remark 9.27 (A Factorization of the System Matrix, Schur Complement Matrix) Let in (9.1) $D = B^T$ and $C = 0$, then a straightforward calculation shows that

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -BA^{-1}B^T \end{pmatrix} \begin{pmatrix} I & A^{-1}B^T \\ 0 & I \end{pmatrix}. \tag{9.8}$$

The matrix

$$S = -BA^{-1}B^T \tag{9.9}$$

is called Schur complement matrix or pressure Schur complement matrix. This matrix is generally not explicitly available. Since it contains the factor A^{-1}, it is usually not a sparse matrix. □

Example 9.28 (Semi-implicit Method for Pressure-Linked Equations (SIMPLE)) SIMPLE is a classical preconditioner that was introduced in Patankar (1980). Its principal approach is as follows:

- An approximation of the pressure is assumed to be known, e.g., from the previous iteration.
- Then, the velocity is computed from the momentum equation, see (9.14). This velocity will usually not satisfy the discrete continuity equation since the used pressure is only an approximation of the pressure that corresponds to the new velocity.

- Finally, the velocity and the pressure are corrected such that the discrete continuity equation is satisfied, see (9.15) and (9.16).

From the point of view of linear algebra, SIMPLE relies on the factorization

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} A & 0 \\ B & -BA^{-1}B^T \end{pmatrix} \begin{pmatrix} I & A^{-1}B^T \\ 0 & I \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{f_p} \end{pmatrix}, \tag{9.10}$$

where the first and second factor in (9.8) are multiplied. The factorization (9.10) is of the form of a block LU factorization. The SIMPLE method is obtained by the approximation

$$A^{-1} \approx (\mathrm{diag}(A))^{-1} = A_{\mathrm{diag}}^{-1}.$$

The approximation of the Schur complement matrix is denoted by

$$S_{\mathrm{d}} = -BA_{\mathrm{diag}}^{-1}B^T. \tag{9.11}$$

Since B^T represents a discrete gradient operator and B a discrete divergence operator, the matrix S_{d} can be thought of representing a discrete diffusion operator with non-constant diffusion. The non-constant diffusion is given by A_{diag}^{-1}. Thus, one iteration of SIMPLE solves the following system

$$\begin{pmatrix} A & 0 \\ B & S_{\mathrm{d}} \end{pmatrix} \begin{pmatrix} I & A_{\mathrm{diag}}^{-1}B^T \\ 0 & I \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{f_p} \end{pmatrix}. \tag{9.12}$$

In practice, one can compute with one step of the iteration directly the new iterate or one can compute just updates of the old iterate. Here, the version with the updates will be presented. Given the iterates at iteration m, then the ansatz for the next iterates is

$$\begin{pmatrix} \underline{u}^{(m+1)} \\ \underline{p}^{(m+1)} \end{pmatrix} = \begin{pmatrix} \underline{u}^{(m)} \\ \underline{p}^{(m)} \end{pmatrix} + \omega \begin{pmatrix} \underline{\delta u} \\ \underline{\delta p} \end{pmatrix}, \tag{9.13}$$

where $\omega \in (0, 1]$ is some damping factor. Inserting this ansatz in (9.10) gives

$$\omega \begin{pmatrix} A & 0 \\ B & -BA^{-1}B^T \end{pmatrix} \begin{pmatrix} I & A^{-1}B^T \\ 0 & I \end{pmatrix} \begin{pmatrix} \underline{\delta u} \\ \underline{\delta p} \end{pmatrix} = \begin{pmatrix} \underline{f} \\ \underline{f_p} \end{pmatrix} - \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{u}^{(m)} \\ \underline{p}^{(m)} \end{pmatrix} = \begin{pmatrix} \underline{r}_{\mathrm{u}} \\ \underline{r}_{\mathrm{p}} \end{pmatrix}.$$

Now, SIMPLE proceeds as follows:

- Solve

$$\begin{pmatrix} A & 0 \\ B & S_{\mathrm{d}} \end{pmatrix} \begin{pmatrix} \underline{\delta u}^* \\ \underline{\delta p}^* \end{pmatrix} = \begin{pmatrix} \underline{r}_{\mathrm{u}} \\ \underline{r}_{\mathrm{p}} \end{pmatrix},$$

  i.e., one has to solve the first

$$A\underline{\delta u}^* = \underline{r}_{\mathrm{u}} \tag{9.14}$$

  and then

$$S_{\mathrm{d}}\underline{\delta p}^* = \underline{r}_{\mathrm{p}} - B\underline{\delta u}^*. \tag{9.15}$$

- Solve

$$\begin{pmatrix} I & A_{\mathrm{diag}}^{-1}B^T \\ 0 & I \end{pmatrix} \begin{pmatrix} \underline{\delta u} \\ \underline{\delta p} \end{pmatrix} = \begin{pmatrix} \underline{\delta u}^* \\ \underline{\delta p}^* \end{pmatrix},$$

  which is nothing else than to set first

$$\underline{\delta p} = \underline{\delta p}^* \tag{9.16}$$

  and then to set

$$\underline{\delta u} = \underline{\delta u}^* - A_{\mathrm{diag}}^{-1}B^T\underline{\delta p}.$$

- Compute the new iterate with (9.13).

SIMPLE can be used as left and as right preconditioner. The application as left preconditioner gives, using (9.12),

$$\begin{aligned} \left(\begin{pmatrix} A & 0 \\ B & S_{\mathrm{d}} \end{pmatrix} \begin{pmatrix} I & A_{\mathrm{diag}}^{-1}B^T \\ 0 & I \end{pmatrix}\right)^{-1} \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} &= \begin{pmatrix} I & -A_{\mathrm{diag}}^{-1}B^T \\ 0 & I \end{pmatrix} \begin{pmatrix} A^{-1} & 0 \\ -S_{\mathrm{d}}^{-1}BA^{-1} & S_{\mathrm{d}}^{-1} \end{pmatrix} \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \\ &= \begin{pmatrix} I & A^{-1}B^T + A_{\mathrm{diag}}^{-1}B^TS_{\mathrm{d}}^{-1}BA^{-1}B^T \\ 0 & -S_{\mathrm{d}}^{-1}BA^{-1}B^T \end{pmatrix}, \end{aligned}$$

which is an approximation of the identity because (9.11) is an approximation of (9.9). Likewise, one finds for using SIMPLE as right preconditioner

$$
\begin{aligned}
&\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \left(\begin{pmatrix} A & 0 \\ B & S_\mathrm{d} \end{pmatrix} \begin{pmatrix} I & A_\mathrm{diag}^{-1} B^T \\ 0 & I \end{pmatrix} \right)^{-1} \\
&= \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} I & -A_\mathrm{diag}^{-1} B^T \\ 0 & I \end{pmatrix} \begin{pmatrix} A^{-1} & 0 \\ -S_\mathrm{d}^{-1} B A^{-1} & S_\mathrm{d}^{-1} \end{pmatrix} \\
&= \begin{pmatrix} I + A A_\mathrm{diag}^{-1} B^T S_\mathrm{d}^{-1} B A^{-1} - B^T S_\mathrm{d}^{-1} B A^{-1} & -A A_\mathrm{diag}^{-1} B^T S_\mathrm{d}^{-1} + B^T S_\mathrm{d}^{-1} \\ 0 & I \end{pmatrix},
\end{aligned}
$$

which is an approximation of the identity if A_diag^{-1} is an approximation of A^{-1}. □

Remark 9.29 (Concerning SIMPLE) SIMPLE is easily to implement, which makes it attractive. It relies on the already assembled matrix blocks. Only the approximation S_d of the Schur complement matrix given in (9.11) has to be computed. This matrix couples pressure degrees of freedom that are usually not coupled in finite element approximations of the diffusion operator, but S_d is still a sparse matrix.

The efficiency of SIMPLE depends on how good A^{-1} is approximated by A_diag^{-1}. It is known, e.g., from Elman et al. (2008b), that the efficiency is bad for convection-dominated problems since the diagonal of A does not contain sufficient information about the convective term. In addition, it can be observed that the number of iterations increases with mesh refinement, e.g., see Schönknecht (2015). □

Example 9.30 (Least Squares Commutator (LSC) Preconditioner) The inefficient behavior of SIMPLE comes from the fact that the approximation (9.11) of the Schur complement is usually bad for convection-dominated problems. In Elman et al. (2006), a preconditioner is proposed that contains information about the convection, the so-called Least Squares Commutator (LSC) preconditioner.

Starting point is the multiplication of the second and third factor of (9.8) yielding

$$
\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix} \begin{pmatrix} A & B^T \\ 0 & S \end{pmatrix}.
$$

Taking the inverse of the last matrix gives

$$
\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} A & B^T \\ 0 & S \end{pmatrix}^{-1} = \begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix}. \tag{9.17}
$$

This representation leads to the idea that

$$
\begin{pmatrix} A & B^T \\ 0 & S \end{pmatrix}^{-1} \tag{9.18}
$$

might be a good right-oriented preconditioner since the matrix on the right-hand side of (9.17) is a triangular matrix. In order to apply an approximation of this preconditioner, the Schur complement matrix S has to be approximated.

The approximation of the Schur complement is based on the construction of the approximation

$$A\left(D_{\mathrm{lsc}}^{-1}B^T\right) \approx B^T A_{\mathrm{pres}}, \tag{9.19}$$

where A_{pres} represents a discretization of a convection-diffusion operator for the pressure. The diagonal matrix D_{lsc} with positive diagonal entries is a scaling matrix. Taking for the moment D_{lsc} to be the identity, then the left-hand side of (9.19) represents the discretization of a convection-diffusion operator for the velocity applied to a discrete gradient and the right-hand side the discrete gradient applied to a convection-diffusion operator for the pressure, compare Elman et al. (2014, Sect. 9.2.3). If the equal sign would hold in (9.19), there would be a commutation of a convection-diffusion operator and a gradient operator.

Multiplying (9.19) with $-BA^{-1}$ from left and A_{pres}^{-1} from right gives, compare (9.9),

$$S = -BA^{-1}B^T \approx -BD_{\mathrm{lsc}}^{-1}B^T A_{\mathrm{pres}}^{-1} = S_{\mathrm{lsc}}. \tag{9.20}$$

Now, A_{pres} is determined by minimizing the commutation error (9.19) column-by-column in a weighted norm. Denoting the k-th column of a matrix by $[\cdot]_k$, the minimization problem is of the form

$$\min_{\left[A_{\mathrm{pres}}\right]_k} \left\| \left[AD_{\mathrm{lsc}}^{-1}B^T\right]_k - B^T \left[A_{\mathrm{pres}}\right]_k \right\|_{D_{\mathrm{lsc}}^{-1}}^2, \quad k = 1, \ldots, N_p, \tag{9.21}$$

where the norm is given by the vector product

$$\left(\underline{x}, \underline{y}\right)_{D_{\mathrm{lsc}}^{-1}} = \underline{x}^T D_{\mathrm{lsc}}^{-1} \underline{y}.$$

Thus, the normal equation corresponding to the least squares problem (9.21) is given by

$$BD_{\mathrm{lsc}}^{-1}\left(\left[A\left(D_{\mathrm{lsc}}^{-1}B^T\right)\right]_k - B^T\left[A_{\mathrm{pres}}\right]_k\right) = 0.$$

It follows that

$$\left[A_{\mathrm{pres}}\right]_k = \left(BD_{\mathrm{lsc}}^{-1}B^T\right)^{-1}\left(BD_{\mathrm{lsc}}^{-1}\left[AD_{\mathrm{lsc}}^{-1}B^T\right)\right]_k$$

and thus

$$A_{\mathrm{pres}} = \left(BD_{\mathrm{lsc}}^{-1}B^T\right)^{-1}\left(BD_{\mathrm{lsc}}^{-1}AD_{\mathrm{lsc}}^{-1}B^T\right).$$

Inserting this expression in (9.20) gives the approximation of the Schur complement matrix

$$S_{\text{lsc}} = -BD_{\text{lsc}}^{-1}B^T \left(BD_{\text{lsc}}^{-1}AD_{\text{lsc}}^{-1}B^T\right)^{-1} BD_{\text{lsc}}^{-1}B^T. \tag{9.22}$$

In Elman et al. (2006), it is proposed to use the diagonal of the velocity mass matrix, see (7.87), as D_{lsc}.

The application of the LSC preconditioner requires to solve as preconditioning step a problem of the form

$$\begin{pmatrix} A & B^T \\ 0 & S_{\text{lsc}} \end{pmatrix} \begin{pmatrix} \underline{v} \\ \underline{q} \end{pmatrix} = \begin{pmatrix} \underline{b}_v \\ \underline{b}_q \end{pmatrix}$$

for a given vector $\left(\underline{b}_v, \underline{b}_q\right)^T$. In the first step, one solves

$$S_{\text{lsc}}\underline{q} = \underline{b}_q,$$

which requires the solution of two discrete diffusion problems for the pressure with the same matrix $BD_{\text{lsc}}^{-1}B^T$ since

$$S_{\text{lsc}}^{-1} = -\left(BD_{\text{lsc}}^{-1}B^T\right)^{-1}\left(BD_{\text{lsc}}^{-1}AD_{\text{lsc}}^{-1}B^T\right)\left(BD_{\text{lsc}}^{-1}B^T\right)^{-1}. \tag{9.23}$$

After having computed $\underline{q}$, one finds $\underline{v}$ by solving

$$A\underline{v} = \underline{b}_v - B^T\underline{q},$$

which requires the solution of a problem for the velocity unknowns. □

Remark 9.31 (Boundary-Corrected LSC Preconditioner) The LSC preconditioner derived in Example 9.30 does not account for the concrete boundary condition of the considered problem in the construction of A_{pres}. However, it was observed in Elman and Tuminaro (2009) that modifications for some kinds of boundary conditions might improve the efficiency of this preconditioner considerably.

A careful analysis of a one-dimensional model problem in Elman and Tuminaro (2009) showed that the commutator error does not vanish if a commutator of the form (9.19) is considered. Instead, a commutator of the form $BA \approx A_{\text{pres}}B$ can be studied. For this form, the commutator error in the one-dimensional model problem vanishes if for the definition of A_{pres} Robin boundary conditions are used at the inflow and Dirichlet boundary conditions at the outflow.

Investigating rectangular domains whose boundaries are parallel to the coordinate axes, it was shown that it is not possible to find boundary conditions such that all commutator errors vanish at the boundary. An analysis of different boundary conditions and the resulting commutator errors led finally to the proposal that the

LSC preconditioner should be augmented with a diagonal weighting matrix W_{lsc} with

$$(W_{\mathrm{lsc}})_{ii} = \begin{cases} \varepsilon, & \text{if the velocity degree of freedom with index } i \text{ is} \\ & \text{associated with a basis function} (\phi_i, 0)^T \text{ and} \\ & b_{ji} \neq 0 \text{ for some pressure degree of freedom } j \text{ on a} \\ & \text{horizontal boundary } \Gamma_{\mathrm{diri}}, \\ \varepsilon, & \text{if the velocity degree of freedom with index } i \text{is} \\ & \text{associated with a basis function } (0, \phi_i)^T \text{ and} \\ & b_{ji} \neq 0 \text{ for some pressure degree of freedom } j \text{ on a} \\ & \text{vertical boundary } \Gamma_{\mathrm{diri}}, \\ 1, & \text{else.} \end{cases} \tag{9.24}$$

Minimizing the commutator error in the least squares sense similarly as in the derivation of the LSC preconditioner, see Elman et al. (2014, Sect. 9.2.4) for details, gives finally the so-called boundary-corrected LSC preconditioner

$$S_{\text{bdr-lsc}} = -BH^{-1}B^T \left(BD_{\mathrm{lsc}}^{-1}AH^{-1}B^T\right)^{-1} BD_{\mathrm{lsc}}^{-1}B^T$$

with $H = W_{\mathrm{lsc}}^{-1/2} D_{\mathrm{lsc}} W_{\mathrm{lsc}}^{-1/2}$ and D_{lsc} is the diagonal of the velocity mass matrix.

For the weight in (9.24), it is proposed to use $\varepsilon = 0.1$, see Elman et al. (2014, Sect. 9.2.4). It is noted in Elman et al. (2014, p. 379) that the construction of a boundary-corrected LSC preconditioner can be generalized to domains that are not aligned with the coordinate axes and to domains in three dimensions. □

Remark 9.32 (LSC Preconditioner for Inf-Sup Stabilized Discretizations) LSC preconditioners can be constructed also for pairs of finite element spaces that do not satisfy the discrete inf-sup condition (3.51) and where the discretization contains some stabilization with respect to the violation of this condition. So-called stabilized LSC preconditioners for lowest order pairs of spaces, e.g., P_1/P_0 or P_1/P_1, were derived in Elman et al. (2008a). □

Remark 9.33 (Further Preconditioners Involving an Approximation of the Pressure Schur Complement) Three preconditioners needing an approximation of the pressure Schur complement were compared in Olshanskii and Vassilevski (2007). One of them was the LSC preconditioner described in Example 9.30. The other ones were:

- The pressure convection-diffusion (PCD) preconditioner from Kay et al. (2002), see also Elman et al. (2014, Sect. 9.2.1). This preconditioner involves approximations to a convection-diffusion operator and a Laplace operator in the discrete pressure space Q^h. These approximations have to be assembled additionally to the other matrices and they have to take into account the boundary condition of the considered problem.

- Based on investigations of the LSC preconditioner for the mesh width $h \to 0$, a preconditioner is proposed that commutes the application of the discrete scaled pressure Laplacian and the discrete divergence operator, i.e., the first factor of (9.23) with B and the last factor of (9.23) with B^T. This commutation gives rise to the solution of a velocity Poisson problem instead of a pressure Poisson problem.

□

Remark 9.34 (Numerical Experience with LSC Preconditioners) Numerical investigations in Olshanskii and Vassilevski (2007) considered the LSC preconditioner and the preconditioners described in Remark 9.33 in BiCGStab and GMRES. It was shown that all of these preconditioners worked satisfactorily if the viscosity ν was large or of modest size, showing only a mild dependency on ν. The efficiency became notably worse for small ν.

Results of a comparison of the LSC preconditioner with the augmented Lagrangian-based preconditioner are described in Remark 9.36. □

Remark 9.35 (An Augmented Lagrangian-Based Preconditioner) This kind of preconditioner was introduced in Benzi and Olshanskii (2006) and analyzed in Benzi and Wang (2011).

Consider (9.1) with $C = 0$ and $D = B^T$, then the linear saddle point problem (9.1) is equivalent to

$$
\begin{aligned}
\begin{pmatrix} A + \gamma B^T W^{-1} B & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} &= \begin{pmatrix} \underline{f} + \gamma B^T W^{-1} B\, B^T g \\ \underline{f_p} \end{pmatrix} \\
&\Longleftrightarrow \\
\begin{pmatrix} A_\gamma & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} &= \begin{pmatrix} \underline{f}_\gamma \\ \underline{f_p} \end{pmatrix},
\end{aligned}
\tag{9.25}
$$

where W is an arbitrary symmetric positive definite matrix and $\gamma > 0$ is a parameter. It was shown in Benzi and Olshanskii (2006) that a theoretical good choice for the matrix is the pressure mass matrix

$$
\left(M_{\mathrm{p}}\right)_{ij} = \left(\psi_j^h, \psi_i^h\right), \quad i = 1, \ldots, N_p.
$$

In practice, its diagonal $W = \mathrm{diag}(M_{\mathrm{p}})$ is used.

The term $\gamma B^T W^{-1} B$ is the so-called augmented Lagrangian term. Since B represents the negative of a discrete divergence operator, (4.92), and B^T a discrete gradient, the augmented Lagrangian term represents a scaled strong grad-div stabilization term, compare Remark 4.117.

The main goal of the augmented Lagrangian-based preconditioner consists in constructing an upper triangular block matrix of the form

$$\begin{pmatrix} \hat{A}_\gamma & B^T \\ 0 & \hat{S}_\gamma \end{pmatrix}, \tag{9.26}$$

where $\hat{A}_\gamma$ is also an upper triangular block matrix. Consider for simplicity of presentation the two-dimensional case and

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \quad B = (B_1, B_2).$$

A straightforward calculation gives

$$A_\gamma = \begin{pmatrix} A_1 + \gamma B_1^T W^{-1} B_1 & \gamma B_1^T W^{-1} B_2 \\ \gamma B_2^T W^{-1} B_1 & A_2 \gamma B_2^T W^{-1} B_2 \end{pmatrix} = \begin{pmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{pmatrix}.$$

Now, the approximation

$$\hat{A}_\gamma = \begin{pmatrix} \hat{A}_{11} & \hat{A}_{12} \\ 0 & \hat{A}_{22} \end{pmatrix}$$

is used. From the analysis performed in Benzi and Wang (2011), it follows that one can choose $\hat{S}_\gamma^{-1} = \gamma W^{-1}$. The inverse of (9.26) can be written in the form

$$\begin{pmatrix} \hat{A}_\gamma^{-1} & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & B^T \\ 0 & -I \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & -\hat{S}_\gamma \end{pmatrix}. \tag{9.27}$$

It follows from (9.27) that one preconditioning step consists of the solution of one system with the matrix $\hat{S}_\gamma$, which is in practice a diagonal matrix, and one system for each component of the velocity with the matrices $\hat{A}_{22}$ and $\hat{A}_{11}$, respectively. A pressure Schur complement system does not appear in this method.

This approach can be extended in a straightforward way to matrices A with a full block structure and to three-dimensional problems.

The crucial issue for the efficiency of the augmented Lagrangian-based preconditioner is the choice of γ. It was found in Benzi and Wang (2011), based on a Fourier analysis and after having applied several simplifications, that the optimal γ is the solution of a non-convex and non-smooth optimization problem. Since the solution of this optimization problem is too complicated, it was proposed in Benzi and Wang (2011) that a norm of the vector-valued function to be minimized should be computed for a large set of parameters and the parameter which gives the smallest norm should be used.

It is possible to use the augmented Lagrangian-based preconditioner as left or right preconditioner. Only little differences in the efficiency were reported in Benzi and Wang (2011).

The construction of the augmented Lagrangian-based preconditioner can be extended to stabilized discretizations that do not satisfy the discrete inf-sup condition (3.51), see Benzi and Wang (2011). □

Remark 9.36 (Numerical Experience with the Augmented Lagrangian-Based Preconditioner) The augmented Lagrangian-based preconditioner and the LSC preconditioner were compared for the solution of two-dimensional steady-state problems, actually Oseen problems (5.1) with $c = 0$ and $\boldsymbol{b}$ was an iterate obtained with some Picard iterations. For the heuristic parameter choice of γ described in Remark 9.35, both preconditioners behaved similarly for large values of the viscosity and the augmented Lagrangian-based preconditioner was more efficient for small values of the viscosity. The augmented Lagrangian-based preconditioner was also able to cope with stretched grids and with inexact solutions of the linear system of equations arising in its application (9.27).

In a numerical study for a two-dimensional problem in ur Rehman et al. (2008), the augmented Lagrangian-based preconditioner was more efficient than the LSC preconditioner and the LSC preconditioner was more efficient than the PCD preconditioner. □

Appendix A
Functional Analysis

Remark A.1 (Motivation) The study of the existence and uniqueness of solutions of the Navier–Stokes equations as well as the finite element error analysis requires tools from functional analysis, in particular the use of function spaces, certain inequalities, and imbedding theorems. There will be no difference in the notation for functions spaces for scalar, vector-valued, and tensor-valued functions.

Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be a domain, i.e., Ω is an open set. □

A.1 Metric Spaces, Banach Spaces, and Hilbert Spaces

Definition A.2 (Metric Space) Let $X \neq \emptyset$ be a set. A map $d : X \times X \to \mathbb{R}$ is called a metric on X if for all $x, y, z \in X$ it is

(i) $d(x, y) = 0 \iff x = y$,
(ii) symmetry: $d(x, y) = d(y, x)$,
(iii) triangle inequality: $d(x, y) \le d(x, z) + d(z, y)$.

Then (X, d) is called a metric space. □

Definition A.3 (Isometric Metric Space) Two metric spaces (X_1, d_1) and $X_2, d_2)$ are called isometric, if there is a surjective map $g : X_1 \to X_2$ such that for all $x, y \in X_1$ it is $d_1(x, y) = d_2(g(x), g(y))$. □

Definition A.4 (Cauchy Sequence, Convergent Sequence) Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in a metric space (X, d). It is called a Cauchy sequence if for each $\varepsilon > 0$ there is a $N \in \mathbb{N}$ such that

$$d(x_k, x_l) < \varepsilon \quad \forall\, k, l \ge N.$$

V. John, *Finite Element Methods for Incompressible Flow Problems*, Springer Series in Computational Mathematics 51, DOI 10.1007/978-3-319-45750-5

The sequence $\{x_n\}_{n=1}^{\infty}$ converges to $x \in X$, denoted by $x_n \to x$, if

$$\lim_{n\to\infty} d(x_n, x) = 0.$$

□

Definition A.5 (Complete Metric Space) A metric space (X, d) is called complete, if each Cauchy sequence converges in X. That means, for each Cauchy sequence $\{x_n\}_{n=1}^{\infty}$ there exists an element $x \in X$ such that $x_n \to x$. □

Definition A.6 (Norm, Triangle Inequality, Seminorm, Normed Space) Let X be a linear space over $\mathbb{R}$ (or $\mathbb{C}$). A mapping $\|\cdot\|_X : X \to \mathbb{R}$ is called a norm on X if

(i) definiteness: $\|x\|_X = 0$ if and only if $x = 0$,
(ii) homogeneity: $\|\alpha x\|_X = |\alpha| \, \|x\|_X$ for all $x \in X$, $\alpha \in \mathbb{R}$,
(iii) the triangle inequality holds: $\|x + y\|_X \le \|x\|_X + \|y\|_X$ for all $x, y \in X$.

A mapping from X to $\mathbb{R}$ that satisfies only (ii) and (iii) is called a seminorm on X.

The space $(X, \|\cdot\|_X)$ is called normed space. □

Definition A.7 (Equivalent Norms) Two norms $\|\cdot\|_{X,1}$, $\|\cdot\|_{X,2}$ of a normed space X are called equivalent, if there are two positive constants C_1 and C_2 such that

$$C_1 \, \|x\|_{X,1} \le \|x\|_{X,2} \le C_2 \, \|x\|_{X,1} \quad \forall \, x \in X.$$

□

Remark A.8 (On Norms)

- All norms in finite-dimensional spaces are equivalent.
- A normed space $(X, \|\cdot\|_X)$ becomes a metric space with the induced metric

$$d(x_1, x_2) = \|x_1 - x_2\|_X \,, \quad x_1, x_2 \in X.$$

□

Definition A.9 (Banach Space) A normed space is called complete if it is a complete metric space with the induced metric. A complete normed space is called Banach space. □

Remark A.10 (Compact Set, Precompact Set) A subset Y of a normed space X is called compact if every sequence of elements in Y has a subsequence that converges in the norm of X to an element of Y. The set Y is called precompact if its closure $\overline{Y}$ is compact.

Compact sets are closed and bounded. The reverse statement is only true for finite-dimensional spaces. □

Definition A.11 (Inner Product, Scalar Product) Let X be a linear space over $\mathbb{R}$. A map $(\cdot,\cdot)_X : X \times X \to \mathbb{R}$ is called symmetric sesquilinear form if for all $x, y, z \in X$ and all $\alpha \in \mathbb{R}$ it holds that

(i) symmetry: $(x, y)_X = (y, x)_X$,
(ii) $(\alpha x, y)_X = \alpha(x, y)$,
(iii) $(x, y + z) = (x, y) + (x, z)$.

The symmetric sesquilinear form $(\cdot,\cdot)_X$ is called positive semi-definite if for all $x \in X$ it is $(x, x)_X \geq 0$. A positive semi-definite symmetric sesquilinear form with

$$(x, x)_X = 0 \quad \Longleftrightarrow \quad x = 0$$

is called inner product or scalar product on X. □

Definition A.12 (Induced Norm, Inner Product Space, Hilbert Space) Let $(\cdot,\cdot)_X$ be an inner product on X, then $(X, (\cdot,\cdot)_X)$ is called pre Hilbert space. The inner product induces the norm

$$\|x\|_X = (x, x)_X^{1/2}$$

in X. A complete inner product space is called Hilbert space.

For simplicity of notation, the subscript at the inner product symbol will be neglected if the inner product is clear from the context. □

Lemma A.13 (Cauchy–Schwarz Inequality) *Let $(X, (\cdot,\cdot))$ be an inner product space, then it holds the so-called Cauchy–Schwarz inequality*

$$|(x, y)| \leq \|x\|_X \, \|y\|_X \quad \forall\, x, y \in X. \tag{A.1}$$

Example A.14 (Cauchy–Schwarz Inequality for Sums) Consider $X = \mathbb{R}^n$ with the standard inner product for vectors, then one obtains with the triangle inequality and the Cauchy–Schwarz inequality (A.1)

$$\left|\sum_{i=1}^n x_i y_i\right| \leq \sum_{i=1}^n |x_i|\,|y_i| \leq \left(\sum_{i=1}^n x_i^2\right)^{1/2} \left(\sum_{i=1}^n y_i^2\right)^{1/2}, \tag{A.2}$$

for all $\underline{x} = (x_1, \ldots, x_n)^T, \underline{y} = (y_1, \ldots, y_n)^T \in \mathbb{R}^n$. □

Example A.15 (Hölder Inequality for Sums) The Cauchy–Schwarz inequality (A.2) is a special case of the Hölder inequality

$$\sum_{i=1}^n |a_i b_i| \leq \left(\sum_{i=1}^n |a_i|^p\right)^{1/p} \left(\sum_{i=1}^n |b_i|^q\right)^{1/q}, \quad 1 < p, q < \infty,\ \frac{1}{p} + \frac{1}{q} = 1. \tag{A.3}$$

The following inequalities for sums of non-negative real numbers hold:

$$\sum_{i=1}^{n} a_i \le \left(\sum_{i=1}^{n} a_i^{1/p}\right)^p \le n^{p/q} \sum_{i=1}^{n} a_i, \quad a_i \ge 0,\ p \in (1,\infty),\ \frac{1}{p} + \frac{1}{q} = 1. \tag{A.4}$$

The right inequality of (A.4) is just a consequence of (A.3). □

Definition A.16 (Orthogonal Elements, Orthogonal Complement of a Subspace) Let X be a normed space endowed with an inner product $(\cdot,\cdot)$. Two elements $x, y \in X$ are said to be orthogonal if $(x, y) = 0$.

Let $Y \subset X$ be a subspace of X, then $Y^\perp = \{x \in X : (x, y) = 0 \text{ for all } y \in Y\}$ is the orthogonal complement of Y. □

Lemma A.17 (Orthogonal Complement is Closed Subspace) *Let $W \subset V$ be a subspace of a Hilbert space V. Then, $W^\perp$ is a closed subspace of V.*

Lemma A.18 (Young's Inequality) *Let $a, b \in \mathbb{R}$, then the following inequality is called Young's inequality:*

$$ab \le \frac{t}{p} a^p + \frac{t^{-q/p}}{q} b^q, \quad \frac{1}{p} + \frac{1}{q} = 1, \quad 1 < p, q < \infty, \quad t > 0. \tag{A.5}$$

Proof The proof is based on the strict convexity of the exponential, which follows from the strict positivity of the second derivative. This property reads for $\alpha, \beta \in \mathbb{R}$ and p, q as in (A.5)

$$\exp\left(\frac{\alpha}{p} + \frac{\beta}{q}\right) \le \frac{1}{p}\exp(\alpha) + \frac{1}{q}\exp(\beta).$$

Choosing $\alpha = \ln(ta^p)$ and $\beta = \ln\left(t^{-q/p} b^q\right)$ gives (A.5). ■

Lemma A.19 (Estimate for a Rayleigh Quotient) *Let $A \in \mathbb{R}^{m\times n}$ be a matrix, then it is*

$$\inf_{\underline{x}\in\mathbb{R}^n, \underline{x}\ne\underline{0}} \frac{\underline{x}^T A^T A \underline{x}}{\underline{x}^T \underline{x}} = \lambda_{\min}\left(A^T A\right),$$

where $\lambda_{\min}\left(A^T A\right)$ is the smallest eigenvalue of $A^T A$. The infimum is taken, i.e., it is even a minimum. The quotient on the left-hand side is called Rayleigh quotient.

Proof The matrix $A^T A$ is symmetric and positive semi-definite. Hence, all eigenvalues are non-negative, the (normalized eigenvectors) $\{\underline{\phi}_i\}_{i=1}^n$ form a basis of $\mathbb{R}^n$, and they are mutually orthonormal. Let the eigenvalues be ordered such that

$$0 \le \lambda_{\min}\left(A^T A\right) = \lambda_1 \le \lambda_2 \le \ldots \le \lambda_n.$$

Each vector $\underline{x} \in \mathbb{R}^n$ can be written in the form $\underline{x} = \sum_{i=1}^n x_i \underline{\phi}_i$. Using that the eigenvectors are orthonormal, it follows that $\underline{x}^T \underline{x} = \sum_{i=1}^n x_i^2$ and

$$\underline{x}^T A^T A \underline{x} = \underline{x}^T \sum_{i=1}^n x_i A^T A \underline{\phi}_i = \sum_{j=1}^n \sum_{i=1}^n \lambda_i x_j x_i \underline{\phi}_j \underline{\phi}_i = \sum_{i=1}^n \lambda_i x_i^2 \geq \lambda_{\min}\left(A^T A\right) \sum_{i=1}^n x_i^2.$$

Hence, one gets

$$\inf_{\underline{x} \in \mathbb{R}^n, \underline{x} \neq \underline{0}} \frac{\underline{x}^T A^T A \underline{x}}{\underline{x}^T \underline{x}} \geq \lambda_{\min}\left(A^T A\right).$$

Choosing $\underline{x} = x_1 \underline{\phi}_1$, $x_1 \neq 0$, leads to

$$\frac{\underline{x}^T A^T A \underline{x}}{\underline{x}^T \underline{x}} = \frac{\lambda_1 x_1^2}{x_1^2} = \lambda_1 = \lambda_{\min}\left(A^T A\right),$$

such that the equal sign holds. ■

A.2 Function Spaces

Definition A.20 (Derivatives and Multi-index) A multi-index $\boldsymbol{\alpha}$ is a vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ with $\alpha_i \in \mathbb{N} \cup \{0\}$, $i = 1, \ldots, n$. Derivatives are denoted by

$$D^{\boldsymbol{\alpha}} = \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}}, \quad \text{with} \quad |\boldsymbol{\alpha}| = \sum_{i=1}^n \alpha_i.$$

Low order derivatives are also denoted by subscripts, e.g.,

$$\partial_x u = \frac{\partial u}{\partial x}.$$

□

Definition A.21 (Spaces of Continuously Differentiable Functions $C^m(\Omega)$, $C^m(\overline{\Omega})$, and $C_B^m(\Omega)$) Let $m \in \mathbb{N} \cup \{0\}$, then the space of m-times continuously differentiable functions in Ω is denoted by

$$C^m(\Omega) = \big\{ f \ : \ f \text{ and all its derivatives up to order } m \\ \text{are continuous in } \Omega \big\}.$$

It is

$$C^{\infty}(\Omega) = \bigcap_{m=0}^{\infty} C^m(\Omega).$$

The space $C^m(\overline{\Omega})$ for $m < \infty$ is defined by

$$C^m(\overline{\Omega}) = \big\{f \ : \ f \in C^m(\Omega) \text{ and all derivatives can be extended continuously to } \overline{\Omega}\big\}.$$

One defines

$$C^{\infty}(\overline{\Omega}) = \bigcap_{m=0}^{\infty} C^m(\overline{\Omega}).$$

Finally, the following space is introduced

$$C_B^m(\Omega) = \big\{f \ : \ f \in C^m(\Omega) \text{ and } f \text{ is bounded}\big\}. \tag{A.6}$$

□

Remark A.22 (Spaces of Continuously Differentiable Functions $C^m(\Omega)$, $C^m(\overline{\Omega})$, and $C_B^m(\Omega)$)

- If Ω is bounded, then $C^m(\overline{\Omega})$, equipped with the norm

$$\|f\|_{C^m(\overline{\Omega})} = \sum_{0\le|\boldsymbol{\alpha}|\le m} \max_{\boldsymbol{x}\in\overline{\Omega}} |D^{\boldsymbol{\alpha}} f(\boldsymbol{x})| ,$$

  is a Banach space.
- The space $C_B^m(\Omega)$ becomes a Banach space with the norm

$$\|f\|_{C_B^m(\Omega)} = \max_{0\le|\boldsymbol{\alpha}|\le m} \sup_{\boldsymbol{x}\in\Omega} |D^{\boldsymbol{\alpha}} f(\boldsymbol{x})| .$$

- It is

$$C^m(\overline{\Omega}) \subset C_B^m(\Omega) \subset C^m(\Omega).$$

  Consider, e.g., $\Omega = (0, 1)$ and $f(x) = \sin(1/x)$, then $f \in C_B(\Omega)$ but $f \notin C(\overline{\Omega})$.

□

Definition A.23 (Support) Let $f \in C(\Omega)$, then

$$\text{supp}(f) = \overline{\{\boldsymbol{x} \ : \ f(\boldsymbol{x}) \neq 0\}}$$

is the support of $f(\boldsymbol{x})$. The closure is taken with respect to $\mathbb{R}^d$. A function $f \in C(\Omega)$ is said to have a compact support, if the support of $f(\boldsymbol{x})$ is bounded in $\mathbb{R}^d$ and if $\operatorname{supp}(f) \subset \Omega$. □

Definition A.24 (The Space $C_0^m(\Omega)$) The space $C_0^m(\Omega)$ is given by

$$C_0^m(\Omega) = \{f \ : f \in C^m(\Omega) \text{ and } \operatorname{supp}(f) \text{ is compact in } \Omega\}.$$

In the literature, the space $C_0^\infty(\Omega)$ is often denoted by $\mathcal{D}(\Omega)$.

An important space for the study of the Navier–Stokes equations is

$$C_{0,\mathrm{div}}^\infty(\Omega) = \{\boldsymbol{f} \ : \ \boldsymbol{f} \in C_0^\infty(\Omega), \nabla \cdot \boldsymbol{f} = 0\}. \tag{A.7}$$

□

Definition A.25 (The Spaces $C^{m,\alpha}(\overline{\Omega})$, Spaces of Hölder Continuous Functions) Let $M \in \mathbb{R}^d$, $d \in \{2, 3\}$, be a set and let $\alpha \in (0, 1]$. Then, the constant

$$|f|_{C^{0,\alpha}(M)} = \sup_{x \neq y \in M} \left\{ \frac{|f(\boldsymbol{x}) - f(\boldsymbol{y})|}{|\boldsymbol{x} - \boldsymbol{y}|^\alpha} \right\}$$

is called Hölder coefficient or Hölder constant. For $\alpha = 1$, it is usually called Lipschitz constant.

Let Ω be bounded. For $m \in \mathbb{N} \cup \{0\}$, the following spaces are defined

$$C^{m,\alpha}(\overline{\Omega}) = \left\{f \in C^m(\overline{\Omega}) \ : \ \left|D^{\boldsymbol{\beta}} f\right|_{C^{0,\alpha}(\overline{\Omega})} < \infty, |\boldsymbol{\beta}| = m\right\}.$$

For $m = 0$, these spaces are called spaces of Hölder continuous functions and for $\alpha = 1$, space of Lipschitz continuous functions. □

Remark A.26 (The Spaces $C^{m,\alpha}(\overline{\Omega})$) The spaces $C^{m,\alpha}(\overline{\Omega})$ are Banach spaces if they are equipped with the norm

$$\|f\|_{C^{m,\alpha}(\overline{\Omega})} = \|f\|_{C^m(\overline{\Omega})} + \sum_{|\boldsymbol{\beta}|=m} [D^{\boldsymbol{\beta}} f]_{C^{0,\alpha}(\overline{\Omega})}.$$

□

Definition A.27 (Spaces of (Lebesgue) Integrable Functions $L^p(\Omega)$) The Lebesgue spaces are defined by

$$L^p(\Omega) = \left\{f \ : \ \int_\Omega |f(\boldsymbol{x})|^p \, d\boldsymbol{x} < \infty\right\}, \quad p \in [1, \infty),$$

where the integral is to be understood in the sense of Lebesgue. The space $L^\infty(\Omega)$ is the space of all functions that are bounded for almost all $\boldsymbol{x} \in \Omega$

$$L^\infty(\Omega) = \{f \ : \ |f(\boldsymbol{x})| < \infty \text{ for almost all } \boldsymbol{x} \in \Omega\}.$$

□

Remark A.28 (Lebesgue Spaces)

- The space $L^p(\Omega)$ is a normed vector space with norm

$$\|f\|_{L^p(\Omega)} = \left(\int_\Omega |f(\boldsymbol{x})|^p \, d\boldsymbol{x}\right)^{1/p}, \quad p \in [1, \infty).$$

- An important special case is $L^2(\Omega)$ since this space is a Hilbert space. The inner product $(f, g)_{L^2(\Omega)}$ of $L^2(\Omega)$ and the induced norm are given by

$$(f, g)_{L^2(\Omega)} = \int_\Omega f(\boldsymbol{x})g(\boldsymbol{x}) \, d\boldsymbol{x}, \quad \|f\|_{L^2(\Omega)} = (f, f)^{1/2}_{L^2(\Omega)}.$$

- The space $L^\infty(\Omega)$ becomes a Banach space if it is equipped with the norm

$$\|f\|_{L^\infty(\Omega)} = \operatorname*{ess\,sup}_{\boldsymbol{x}\in\Omega} |f(\boldsymbol{x})|,$$

  where $\operatorname{ess\,sup}_{\boldsymbol{x}\in\Omega}$ is the essential supremum.
- Let $|\Omega| < \infty$ and $1 \le p \le q \le \infty$. If $u \in L^q(\Omega)$, then $u \in L^p(\Omega)$ and

$$\|u\|_{L^p(\Omega)} \le \left(\int_\Omega d\boldsymbol{x}\right)^{1/p-1/q} \|u\|_{L^q(\Omega)}, \tag{A.8}$$

  see Adams (1975, Theorem 2.8).

□

Example A.29 (Cauchy–Schwarz Inequality and Hölder's Inequality) Let $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$ with $p, q \in [1, \infty]$ and $1/p + 1/q = 1$. Then it is $fg \in L^1(\Omega)$ and the Hölder inequality holds

$$\|fg\|_{L^1(\Omega)} \le \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)}. \tag{A.9}$$

For $p = q = 2$, this inequality is called Cauchy–Schwarz inequality

$$\|fg\|_{L^1(\Omega)} \le \|f\|_{L^2(\Omega)} \|g\|_{L^2(\Omega)}. \tag{A.10}$$

□

Definition A.30 (Sobolev Spaces $W^{k,p}(\Omega)$) Let $k \in \mathbb{N}$ and $p \in [1,\infty]$. The Sobolev space $W^{k,p}(\Omega)$ consists of all integrable functions $f : \Omega \to \mathbb{R}$ such that for each multi-index $\boldsymbol{\alpha}$ with $|\boldsymbol{\alpha}| \le k$, the derivative $D^{\boldsymbol{\alpha}} f$ exists in the weak sense and it belongs to $L^p(\Omega)$. □

Remark A.31 (Sobolev Spaces)

- It is $L^p(\Omega) = W^{0,p}(\Omega)$.
- A norm in Sobolev spaces is defined by

$$\|f\|_{W^{k,p}(\Omega)} = \begin{cases} \left(\sum_{|\boldsymbol{\alpha}|\le k} \|D^{\boldsymbol{\alpha}} f\|^p_{L^p(\Omega)}\right)^{1/p} & \text{if } p \in [1,\infty), \\ \sum_{|\boldsymbol{\alpha}|\le k} \operatorname{ess\,sup}_{\boldsymbol{x}\in\Omega} |D^{\boldsymbol{\alpha}} f| & \text{if } p = \infty. \end{cases}$$

  Sobolev spaces equipped with this norm are Banach spaces, e.g., see Evans (2010, p. 262).
- The Sobolev spaces for $p = 2$ are Hilbert spaces. They are often denoted by $W^{m,2}(\Omega) = H^m(\Omega)$ and they are equipped with the inner product

$$(f,g)_{H^k(\Omega)} = \sum_{|\boldsymbol{\alpha}|\le k} (D^{\boldsymbol{\alpha}} f, D^{\boldsymbol{\alpha}} g)_{L^2(\Omega)} .$$

- In particular, the Sobolev spaces of first order are important for the study of the Navier–Stokes equations

$$W^{1,p}(\Omega) = \left\{ f \;:\; \int_\Omega |f(\boldsymbol{x})|^p + |\nabla f(\boldsymbol{x})|^p \, d\boldsymbol{x} < \infty \right\}, \quad p \in [1,\infty),$$

  which are equipped with the norm

$$\|f\|_{W^{1,p}(\Omega)} = \left(\int_\Omega |f(\boldsymbol{x})|^p + |\nabla f(\boldsymbol{x})|^p \, d\boldsymbol{x} \right)^{1/p}, \quad p \in [1,\infty).$$

- The definition of Sobolev spaces can be extended to $k \in \mathbb{R}$, e.g., see Adams (1975).

□

Definition A.32 (Sobolev Spaces $W_0^{k,p}(\Omega)$) The Sobolev spaces $W_0^{k,p}(\Omega)$ are defined by the closure of $C_0^\infty(\Omega)$ in the norm of $W^{k,p}(\Omega)$. □

Remark A.33 (On the Smoothness of the Boundary) The Sobolev imbedding theorem requires that Ω has the so-called cone property or the strong local Lipschitz property. In the case that Ω is bounded, these assumptions reduce to the requirement that Ω has a locally Lipschitz boundary, Adams (1975, p. 67). That means, each point $\boldsymbol{x}$ on the boundary $\partial\Omega$ of Ω has a neighborhood $U_{\boldsymbol{x}}$ such the $\partial\Omega \cap U_{\boldsymbol{x}}$ is the graph of a Lipschitz continuous function. □

Theorem A.34 (Trace Theorem, Lions and Magenes (1972, Theorem 9.4), Galdi (2011, Theorem II.4.1 for $m = 1$)) *Let Ω be a bounded domain with locally Lipschitz boundary $\partial\Omega$. Then, there is a bounded linear operator $T : W^{1,q}(\Omega) \to L^r(\partial\Omega)$, $q \in [1,\infty)$, such that*

(i) $r \in [1, q(d-1)/(d-q)]$ *if* $q < d$ *and* $r \in [1,\infty)$ *else,*
(ii) $Tf = f|_{\partial\Omega}$ *iff* $f \in W^{1,q}(\Omega) \cap C\left(\overline{\Omega}\right)$,
(iii) $\|Tf\|_{L^r(\partial\Omega)} \le C\,\|f\|_{W^{1,q}(\Omega)}$ *for each* $f \in W^{1,q}(\Omega)$, *with the constant* C *depending only on* q *and* Ω.

The mapping

$$H^s(\Omega) \to \prod_{j=0}^{s_0} H^{s-j-1/2}(\partial\Omega), \quad f \mapsto \left\{\frac{\partial^j f}{\partial \boldsymbol{n}^j}, j = 0, 1, \ldots, s_0\right\} \tag{A.11}$$

is continuous, where s_0 is the greatest integer such that $s_0 < s - 1/2$, and $\boldsymbol{n}$ is the outward pointing unit normal vector. The mapping is surjective and there exists a continuous right inverse.

Theorem A.35 (Functions with Vanishing Trace, Galdi (2011, Theorem II.4.2), Evans (2010, p. 273)) *Let the assumptions of Theorem A.34 be given. Then $f \in W_0^{1,p}(\Omega)$ if and only if $Tf = 0$ on $\partial\Omega$.*

Theorem A.36 (Poincaré's Inequality, Poincaré–Friedrichs' Inequality, Galdi (2011, Theorem II.5.1), Gilbarg and Trudinger (1983, p. 164)) *Let $f \in W_0^{1,p}(\Omega)$, then*

$$\|f\|_{L^p(\Omega)} \le \left(\frac{|\Omega|}{\omega_d}\right)^{1/d} \|\nabla f\|_{L^p(\Omega)} = C_{\mathrm{PF}}\,\|\nabla f\|_{L^p(\Omega)} \quad p \in [1,\infty), \tag{A.12}$$

where ω_d is the volume of the unit ball in $\mathbb{R}^d$.

Remark A.37 (Poincaré's Inequality) Poincaré's inequality (A.12) holds also for functions $v \in H^1(\Omega)$ with $v = 0$ on $\Gamma_0 \subset \Gamma$ with $|\Gamma_0| > 0$.

Poincaré's inequality stays valid for vector-valued functions $\boldsymbol{v}$ if Ω is bounded with a locally Lipschitz boundary, $\boldsymbol{v} \in W^{1,q}(\Omega)$, $1 \le q < \infty$, and $\boldsymbol{v} \cdot \boldsymbol{n} = 0$ on $\partial\Omega$, see Galdi (1994, Sect. II.4). □

Theorem A.38 (Density of Continuous Functions in Sobolev Spaces, Gilbarg and Trudinger (1983, p. 154)) *The subspace $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$.*

Remark A.39 (Density of Continuous Functions in Sobolev Spaces) For $C^\infty(\overline{\Omega})$ to be dense in $W^{k,p}(\Omega)$, one needs some smoothness assumptions on the boundary $\partial\Omega$, like $\partial\Omega$ is C^1 or the so-called segment property, e.g., see Gilbarg and Trudinger (1983, p. 155). This segment property follows from the strong local Lipschitz property, see Adams (1975, p. 67). □

Theorem A.40 (Interpolation Theorem for Sobolev Spaces, Adams (1975, Theorem 4.17)) *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a locally Lipschitz boundary and let $p \in [1, \infty)$. Then there exists a constant $C(m, p, \Omega)$ such that for $0 \le j \le m$ and any $u \in W^{m,p}(\Omega)$*

$$\|u\|_{W^{j,p}(\Omega)} \le C(m, p, \Omega)\, \|u\|_{W^{m,p}(\Omega)}^{j/m}\, \|u\|_{L^p(\Omega)}^{(m-j)/m}. \tag{A.13}$$

In addition, (A.13) is valid for all $u \in W_0^{m,p}(\Omega)$ with a constant $C(m, p, d)$ independent of Ω.

Remark A.41 (Imbedding Theorems) Imbedding theorems for Sobolev spaces are used frequently in the analysis of partial differential equations. The imbedding theorems state that all functions belonging to a certain space do belong also to another space and that the norm of the functions in the larger space can be estimated by the norm in the smaller space. Let V be a Banach space such that an imbedding $W^{m,p}(\Omega) \to V$ holds. Then, there is a constant C depending on Ω such that

$$\|v\|_V \le C\, \|v\|_{W^{m,p}(\Omega)}$$

for all functions $v \in W^{m,p}(\Omega)$. The validity of imbeddings depends on the dimension d of the domain Ω. The larger the dimension, the less imbeddings are valid, compare Example A.44. □

Theorem A.42 (The Sobolev Imbedding Theorem, Adams (1975, Theorem 5.4, Remark 5.5 (6), Theorem 6.2)) *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a locally Lipschitz boundary. Let j and m be non-negative integers and let p satisfy $1 \le p < \infty$.*

(i) Let $mp < d$, then the imbedding

$$W^{j+m,p}(\Omega) \to W^{j,q}(\Omega), \quad 1 \le q \le \frac{dp}{d - mp} \tag{A.14}$$

holds. In particular, it is

$$W^{m,p}(\Omega) \to L^q(\Omega), \quad 1 \le q \le \frac{dp}{d - mp}. \tag{A.15}$$

(ii) Suppose $mp = d$. Then the imbedding

$$W^{m,p}(\Omega) \to L^q(\Omega), \quad 1 \le q < \infty \tag{A.16}$$

is valid. If in addition $p = 1$, then this imbedding holds also for $q = \infty$

$$W^{d,1}(\Omega) \to L^\infty(\Omega). \tag{A.17}$$

and even

$$W^{d,1}(\Omega) \to C_B(\Omega),$$

see (A.6) *for the definition of latter space.*

(iii) Suppose that $mp > d$*, then the imbedding*

$$W^{m,p}(\Omega) \to C_B(\Omega) \tag{A.18}$$

holds.

(iv) Suppose $mp > d > (m-1)p$*, then*

$$W^{j+m,p}(\Omega) \to C^{j,\lambda}\left(\overline{\Omega}\right) \quad \text{for} \quad 0 < \lambda \le m - \frac{d}{p}. \tag{A.19}$$

(v) Suppose $d = (m-1)p$*, then*

$$W^{j+m,p}(\Omega) \to C^{j,\lambda}\left(\overline{\Omega}\right) \quad \text{for} \quad 0 < \lambda < 1.$$

This imbedding holds for $\lambda = 1$ *if* $p = 1$ *and* $d = m - 1$*.*

(vi) All imbeddings are true for arbitrary domains provided the W spaces undergoing the imbedding are replaced with the corresponding W_0 *spaces.*

(vii) Rellich–Kondrachov theorem: The imbeddings (A.14)–(A.16) *are compact with the conditions on* Ω *stated at the beginning of the theorem, i.e., the imbedding operator is compact, see Definition A.63.*

Remark A.43 (Spaces of Continuous Functions in $\overline{\Omega}$*)* Since the compact imbedding

$$C^{j,\lambda}\left(\overline{\Omega}\right) \to C^{j}\left(\overline{\Omega}\right) \quad j \ge 0, 0 < \lambda \le 1,$$

holds for bounded domains, Adams (1975, Theorem 1.31), one can derive from Theorem A.42, cases (iv) and (v), also imbeddings for $C^j\left(\overline{\Omega}\right)$: if $mp > d \ge (m-1)p$, $p \in [1, \infty)$, then

$$W^{j+m,p}(\Omega) \to C^{j}\left(\overline{\Omega}\right). \tag{A.20}$$

□

Example A.44 (Important Sobolev Imbeddings) Let $d = 2$. Then, it follows from (A.16) that

$$H^1(\Omega) = W^{1,2}(\Omega) \to L^q(\Omega), \quad q \in [1, \infty). \tag{A.21}$$

For $d = 3$, one gets with (A.15) that

$$H^1(\Omega) = W^{1,2}(\Omega) \to L^q(\Omega), \quad q \in [1, 6]. \tag{A.22}$$

□

Remark A.45 (Spaces of Functions Defined in Space-Time Domains) Let X be any normed space introduced above that is equipped with the norm $\|\cdot\|_X$ and let (t_0, t_1) be a time interval. Then, the following function space on the space-time domain can be defined

$$L^p(t_0, t_1; X) = \left\{ f(t, \boldsymbol{x}) \ : \ \int_{t_0}^{t_1} \|f\|_X^p(\tau)\, d\tau < \infty \right\}, \quad p \in [1, \infty).$$

The norm of $L^p(t_0, t_1; X)$ is

$$\|f\|_{L^p(t_0,t_1;X)} = \left(\int_{t_0}^{t_1} \|f\|_X^p(\tau)\, d\tau \right)^{1/p}, \quad p \in [1, \infty).$$

The modifications for $p = \infty$ are the same as for the Lebesgue spaces. □

A.3 Some Definitions, Statements, and Theorems

Remark A.46 (Convolution) The convolution of two scalar functions f and g is defined by

$$(f * g)(y) = \int_{\mathbb{R}} f(y - x)\, g(x)\ dx = \int_{\mathbb{R}} f(x)\, g(y - x)\ dx = (g * f)(y),$$

provided that the integrals exist for almost all $y \in \mathbb{R}$. □

Remark A.47 (Fourier Transform: Definition and Some Properties) The Fourier transform of a scalar function f is defined by

$$\mathcal{F}(f)(y) = \int_{\mathbb{R}} f(x)\, e^{-ixy}\, dx \tag{A.23}$$

and the inverse Fourier transform of $F(y)$ by

$$\mathcal{F}^{-1}(F)(x) = \frac{1}{2\pi} \int_{\mathbb{R}} F(y)\, e^{ixy}\, dy. \tag{A.24}$$

It holds

$$\mathcal{F}(f * g) = \mathcal{F}(f)\,\mathcal{F}(g)\,, \quad \mathcal{F}(fg) = \mathcal{F}(f) * \mathcal{F}(g)\,. \tag{A.25}$$

If f is differentiable and $\lim_{|x|\to\infty} f(x) = 0$, integration by parts yields

$$y\mathcal{F}(f)(y) = -i\mathcal{F}\left(f'\right)(y)\,.$$

This formula implies the relations

$$\|\boldsymbol{y}\|_2^2\,\mathcal{F}(f) = -\mathcal{F}(\Delta f)\,, \tag{A.26}$$

$$\frac{1}{\|\boldsymbol{y}\|_2^2}\mathcal{F}(f) = -\mathcal{F}\left(\Delta^{-1}(f)\right), \tag{A.27}$$

$$\frac{1}{1 + c\,\|\boldsymbol{y}\|_2^2}\mathcal{F}(f) = \mathcal{F}\left((I - c\Delta)^{-1}(f)\right). \tag{A.28}$$

The $L^r(\Omega)$ norm of $f * g$, $1 \le r \le \infty$, can be estimated by Young's inequality for convolutions (sometimes also called Hölder's inequality for convolutions), e.g., see Hörmander (1990, Sect. IV.4.5). Let $1 \le p, q \le \infty$, $p^{-1} + q^{-1} \ge 1$, and $r^{-1} = p^{-1} + q^{-1} - 1$. For $f \in L^p\left(\mathbb{R}^d\right)$ and $g \in L^q\left(\mathbb{R}^d\right)$, it is $f * g \in L^r\left(\mathbb{R}^d\right)$ and Young's inequality for convolution

$$\|f * g\|_{L^r(\Omega)} \le \|f\|_{L^p(\Omega)}\,\|g\|_{L^q(\Omega)} \tag{A.29}$$

holds. If $p^{-1} + q^{-1} - 1 = 0$, then $f * g \in L^r\left(\mathbb{R}^d\right)$ is continuous and bounded. □

Definition A.48 (Absolutely Continuous Function) Let $I \subset \mathbb{R}$ be an interval. A function $f : I \to \mathbb{R}$ is called absolutely continuous if for every $\varepsilon > 0$ there is a $\delta > 0$ such that whenever a finite sequence of pairwise disjoint sub-intervals (x_k, y_k) of I satisfies $\sum_k (y_k - x_k) < \delta$, then $\sum_k |f(y_k) - f(x_k)| < \varepsilon$. □

Remark A.49 (Absolutely Continuous Functions) Absolute continuity of a function is a stronger condition than continuity and even uniform continuity. On a compact interval $I = [a, b]$, absolute continuity of a function f is equivalent to the property that this function has a derivative f' almost everywhere, the derivative is Lebesgue integrable, and it holds

$$f(t) = f(a) + \int_a^t f'(\tau)\,d\tau \quad \forall\, t \in [a, b]$$

(fundamental theorem of calculus). □

Theorem A.50 (Local Existence and Uniqueness Theorem of Carathéodory, Carathéodory (1918, Kap. 11), Kamke (1944, p. 34), Filippov (1988, Sect. 1.1)) *Let $f_m(t, y_1, \ldots, y_n)$, $m = 1, \ldots, n$, be defined in*

$$\Omega_T = (t_0, t_0 + T) \times \Omega_y$$

with $\Omega_y = \{\boldsymbol{y} = (y_1, \ldots, y_n)^T : \|\boldsymbol{y} - \boldsymbol{y}_0\|_2 \leq b\}$ for some $b > 0$, let the functions $f_1, \ldots, f_n$ for each fixed system $y_1, \ldots, y_n$ be measurable with respect to t, let for each fixed t the functions $f_1, \ldots, f_n$ be continuous with respect to $y_1, \ldots, y_n$, and let

$$|f_m(t, y_1, \ldots, y_n)| \leq F(t), \quad m = 1, \ldots, n,$$

where $F(t)$ is a Lebesgue integrable function in $(t_0, t_0 + T)$. Then there exists a system of absolutely continuous functions $y_1(t), \ldots, y_n(t)$ that satisfies for all t in some interval $[t_0, t_0 + a]$, $0 < a \leq T$,

$$y_m(t) = y_{0m} + \int_{t_0}^{t} f_m(s, y_1(s), \ldots, y_n(s))\, ds, \quad m = 1, \ldots, n. \tag{A.30}$$

At each point where the term in the integral is continuous, the functions satisfy the ordinary differential equation

$$\frac{d}{dt} y_m(t) = f_m(t, y_1, \ldots, y_n), \quad m = 1, \ldots, n. \tag{A.31}$$

If in addition for any two points $(t, \overline{y}_1, \ldots, \overline{y}_n), (t, \hat{y}_1, \ldots, \hat{y}_n) \in \Omega_T$ the Lipschitz condition

$$|f_m(t, \overline{y}_1, \ldots, \overline{y}_n) - f_m(t, \hat{y}_1(t), \ldots, \hat{y}_n(t))| \leq G(t) \sum_{l=1}^{n} |\overline{y}_l - \hat{y}_l(t)|,$$

$m = 1, \ldots, n$, with a Lebesgue integrable function $G(t)$ is satisfied, then there exists exactly one solution of (A.30) in $[t_0, t_0 + a]$.

Remark A.51 (On Carathéodory's Theorem) The theorem of Carathéodory is an extension of the theorem of Peano to ordinary differential equations of type (A.31) with discontinuous right-hand side. □

Remark A.52 (On Gronwall's Lemma) Gronwall's lemma is an important tool for the analysis and finite element analysis of time-dependent problems. Two versions of this lemma in the continuous setting will be given below, see Emmrich (1999) for complete proofs and a discussion of the differences of these versions. □

Lemma A.53 (Gronwall's Lemma in Integral Form) *Let* $T \in \mathbb{R}^+ \cup \infty$, $f, g, \in L^\infty(0,T)$, *and* $\lambda \in L^1(0,T)$, $\lambda(t) \ge 0$ *for almost all* $t \in [0,T]$. *Then*

$$f(t) \le g(t) + \int_0^t \lambda(s) f(s)\ ds \quad a.e.\ in\ [0,T] \tag{A.32}$$

implies for almost all $t \in [0,T]$ *that*

$$f(t) \le g(t) + \int_0^t \exp\left(\int_s^t \lambda(\tau)\ d\tau\right) \lambda(s)\, g(s)\ ds. \tag{A.33}$$

If $g \in W^{1,1}(0,T)$, *it follows that*

$$f(t) \le \exp\left(\int_0^t \lambda(\tau)\ d\tau\right)\left(g(0) + \int_0^t \exp\left(-\int_0^s \lambda(\tau)\ d\tau\right)\right) g'(s)\ ds.$$

Moreover, if $g(t)$ *is a monotonically increasing continuous function, it holds*

$$f(t) \le \exp\left(\int_0^t \lambda(\tau)\ d\tau\right) g(t)\,. \tag{A.34}$$

Proof For illustration, the derivation of (A.33) and (A.34) will be presented.

(A.33). Let

$$\tilde{f}(t) = \exp\left(-\int_0^t \lambda(\tau)\, d\tau\right) \int_0^t \lambda(s) f(s)\, ds,$$

then one obtains for almost all $t \in [0,T]$ with the product rule, the Leibniz integral rule, (A.32), and $\lambda(t) \ge 0$

$$\begin{aligned} \tilde{f}'(t) &= \exp\left(-\int_0^t \lambda(\tau)\, d\tau\right)\left(-\lambda(t)\int_0^t \lambda(s) f(s)\, ds + \lambda(t) f(t)\right) \\ &\le \exp\left(-\int_0^t \lambda(\tau)\, d\tau\right)\left(\lambda(t)\left(g(t) - f(t)\right) + \lambda(t) f(t)\right) \\ &= \exp\left(-\int_0^t \lambda(\tau)\, d\tau\right) \lambda(t) g(t). \end{aligned}$$

Integration yields, using $\tilde{f}(0) = 0$,

$$\tilde{f}(t) \le \int_0^t \exp\left(-\int_0^s \lambda(\tau)\, d\tau\right) \lambda(s) g(s)\, ds.$$

With (A.32), one obtains

$$\exp\left(-\int_0^t \lambda(\tau)\,d\tau\right)(f(t)-g(t)) \le \exp\left(-\int_0^t \lambda(\tau)\,d\tau\right)\int_0^t \lambda(s)f(s)\ ds$$
$$= \tilde{f}(t) \le \int_0^t \exp\left(-\int_0^s \lambda(\tau)\,d\tau\right)\lambda(s)g(s)\,ds.$$

Multiplying this inequality with $\exp\left(\int_0^t \lambda(\tau)\,d\tau\right)$ and bringing $g(t)$ to the right-hand side gives (A.33).

Equation (A.34). If $g(t)$ is a monotonically increasing continuous function, one gets from (A.33), using that $g(t)$ takes its largest value at the final time and that $\lambda(t) \ge 0$, and applying the fundamental theorem of calculus

$$\begin{aligned}
f(t) &\le g(t)\left(1+\int_0^t \exp\left(\int_s^t \lambda(\tau)\ d\tau\right)\lambda(s)\ ds\right)\\
&= g(t)\left(1+\exp\left(\int_0^t \lambda(\tau)\ d\tau\right)\int_0^t \exp\left(-\int_0^s \lambda(\tau)\ d\tau\right)\lambda(s)\ ds\right)\\
&= g(t)\left(1+\exp\left(\int_0^t \lambda(\tau)\ d\tau\right)\int_0^t \frac{d}{ds}\left(-\exp\left(-\int_0^s \lambda(\tau)\ d\tau\right)\right)\ ds\right)\\
&= g(t)\left(1+\exp\left(\int_0^t \lambda(\tau)\ d\tau\right)\left(-\exp\left(-\int_0^t \lambda(\tau)\ d\tau\right)+1\right)\right)\\
&= \exp\left(\int_0^t \lambda(\tau)\ d\tau\right)g(t).
\end{aligned}$$

■

Lemma A.54 (Gronwall's Lemma in Differential Form) *Let $T \in \mathbb{R}^+ \cup \infty$, $f \in W^{1,1}(0,T)$, and $g, \lambda \in L^1(0,T)$. Then*

$$f'(t) \le g(t) + \lambda(t)f(t) \quad a.e.\ in\ [0,T] \tag{A.35}$$

implies for almost all $t \in [0,T]$

$$f(t) \le \exp\left(\int_0^t \lambda(\tau)\,d\tau\right)f(0) + \int_0^t \exp\left(\int_s^t \lambda(\tau)\,d\tau\right)g(s)\,ds. \tag{A.36}$$

Proof Defining

$$\tilde{f}(t) = \exp\left(-\int_0^t \lambda(\tau)\,d\tau\right)f(t) = \exp(-\Lambda(t))f(t), \tag{A.37}$$

applying the chain rule, the Leibniz integral rule, and (A.35) gives

$$\tilde{f}'(t) = -\Lambda'(t)\exp(-\Lambda(t))f(t) + \exp(-\Lambda(t))f'(t) = \exp(-\Lambda(t))\big(f'(t) - \lambda(t)f(t)\big)$$
$$\leq \exp(-\Lambda(t))g(t).$$

Integration in $(0, t)$ and using (A.37) yields

$$\tilde{f}(t) - \tilde{f}(0) = \exp\left(-\int_0^t \lambda(\tau)\,d\tau\right)f(t) - f(0) \leq \int_0^t \exp(-\Lambda(s))g(s)\,ds.$$

Multiplication with $\exp\left(\int_0^t \lambda(\tau)\,d\tau\right)$ gives (A.36). ■

Lemma A.55 (Variation of Gronwall's Lemma in Differential Form) *Let $T \in \mathbb{R}^+ \cup \infty$, $f \in W^{1,1}(0,T)$, $h, g, \lambda \in L^1(0,T)$, and $h(t), \lambda(t) \geq 0$ a.e. in $(0,T)$. Then,*

$$f'(t) + h(t) \leq g(t) + \lambda(t)f(t) \quad a.e.\ in\ [0,T] \tag{A.38}$$

implies for almost all $t \in [0,T]$

$$f(t) + \int_0^t h(s)\,ds \tag{A.39}$$
$$\leq \exp\left(\int_0^t \lambda(\tau)\,d\tau\right)f(0) + \int_0^t \exp\left(\int_s^t \lambda(\tau)\,d\tau\right)g(s)\,ds.$$

Moreover, if $g(t) \geq 0$ a.e. in $(0,T)$, it holds

$$f(t) + \int_0^t h(s)\,ds \leq \exp\left(\int_0^t \lambda(\tau)\,d\tau\right)\left(f(0) + \int_0^t g(s)\,ds\right). \tag{A.40}$$

Proof From (A.38), it follows a.e. in $[0,T]$ that

$$f'(s) - \lambda(s)f(s) + h(s) \leq g(s).$$

The positivity of the exponential implies

$$\exp\left(-\int_0^s \lambda(\tau)\,d\tau\right)\big(f'(s) - \lambda(s)f(s) + h(s)\big) \leq \exp\left(-\int_0^s \lambda(\tau)\,d\tau\right)g(s).$$

Integration on $(0, t) \subset [0, T]$ gives

$$\exp\left(-\int_0^t \lambda(\tau)\ d\tau\right) f(t) - f(0) + \int_0^t \exp\left(-\int_0^s \lambda(\tau)\ d\tau\right) h(s)\ ds \\ \leq \int_0^t \exp\left(-\int_0^s \lambda(\tau)\ d\tau\right) g(s)\ ds. \quad \text{(A.41)}$$

Using the monotonicity of the exponential yields

$$\exp\left(-\int_0^t \lambda(\tau)\ d\tau\right) \int_0^t h(s)\ ds \leq \int_0^t \exp\left(-\int_0^s \lambda(\tau)\ d\tau\right) h(s)\ ds.$$

Applying this inequality to bound the left-hand side of (A.41) from below and multiplication of the resulting inequality with $\exp\left(\int_0^t \lambda(\tau)\ d\tau\right)$ proves (A.39).

If g is non-negative, one obtains

$$\int_0^t \exp\left(\int_s^t \lambda(\tau)\ d\tau\right) g(s)\ ds \leq \exp\left(\int_0^t \lambda(\tau)\ d\tau\right) \int_0^t g(s)\ ds,$$

from which (A.40) follows. ■

Lemma A.56 (Discrete Gronwall's Lemma, Heywood and Rannacher (1990, Lemma 5.1)) *Let $k, B, a_n, b_n, c_n, \alpha_n$ be non-negative numbers for integers $n \geq 1$ and let the inequality*

$$a_{N+1} + k\sum_{n=1}^{N+1} b_n \leq B + k\sum_{n=1}^{N+1} c_n + k\sum_{n=1}^{N+1} \alpha_n a_n \quad \text{for } N \geq 0 \quad \text{(A.42)}$$

hold. If $k\alpha_n < 1$ for all $n = 1, \ldots, N+1$, then

$$a_{N+1} + k\sum_{n=1}^{N+1} b_n \leq \exp\left(k\sum_{n=1}^{N+1} \frac{\alpha_n}{1 - k\alpha_n}\right)\left(B + k\sum_{n=1}^{N+1} c_n\right) \quad \text{for } N \geq 0. \quad \text{(A.43)}$$

If the inequality

$$a_{N+1} + k\sum_{n=1}^{N+1} b_n \leq B + k\sum_{n=1}^{N+1} c_n + k\sum_{n=1}^{N} \alpha_n a_n \quad \text{for } N \geq 0 \quad \text{(A.44)}$$

is given, then it holds

$$a_{N+1} + k\sum_{n=1}^{N+1} b_n \leq \exp\left(k\sum_{n=1}^{N} \alpha_n\right)\left(B + k\sum_{n=1}^{N+1} c_n\right) \quad \text{for } N \geq 0. \quad \text{(A.45)}$$

Definition A.57 (Weak Convergence and Weak* Convergence, Yosida (1995, p. 120, p. 125)) A sequence $\{x_n\}_{n=1}^{\infty}$ in a normed linear space X is said to be weakly convergent if a finite limit $\lim_{n\to\infty} f(x_n)$ exists for each $f \in X'$, where X' is the (strong) dual space of X. If

$$\lim_{n\to\infty} f(x_n) = f(x) \quad \forall\, f \in X',$$

then $\{x_n\}_{n=1}^{\infty}$ is called to be weakly convergent to x, in notation $x_n \rightharpoonup x$.

A sequence $\{f_n\}_{n=1}^{\infty}$ in the (strong) dual space X' of a linear space X is said to be weakly* convergent if a finite limit $\lim_{n\to\infty} f_n(x)$ exists for every $x \in X$. The sequence is said to converge weakly* to an element $f \in X'$ if

$$\lim_{n\to\infty} f_n(x) = f(x) \quad \forall\, x \in X,$$

in notation $f_n \overset{*}{\rightharpoonup} f$. □

Remark A.58 (On the Weak and Weak* Convergence)

- If the limit x or f exist, then the limit is unique, Yosida (1995, p. 120).
- If X is a reflexive Banach space and if $\{x_n\}_{n=1}^{\infty} \subset X$ is bounded, then there exists a subsequence $\{x_{n_l}\}_{l=1}^{\infty} \subset \{x_n\}_{n=1}^{\infty}$ and an element $x \in X$ such that $x_{n_l} \rightharpoonup x$, see Evans (2010, p. 723).
- If $\{f_n\}_{n=1}^{\infty} \subset X'$ is bounded in the dual X' of X and X is a separable Banach space, then there exists a weakly* convergent subsequence.

□

Definition A.59 (Linear Operator, Range, Kernel) Let X and Y be real Banach spaces. A mapping $A : X \to Y$ is a linear operator if

$$A(\alpha x_1 + \beta x_2) = \alpha A x_1 + \beta A x_2 \quad \forall\, x_1, x_2 \in X, \alpha, \beta \in \mathbb{R}.$$

The range or image of A is given by

$$\text{range}(A) = \{y \in Y : y = Ax \text{ for some } x \in X\}.$$

The kernel or the null space of A is defined by

$$\ker(A) = \{x \in X : Ax = 0\}.$$

□

Definition A.60 (Bounded Operator, Continuous Operator) An operator $A : X \to Y$, X, Y Banach spaces, is bounded if

$$\|A\| = \sup_{x\in X} \frac{\|Ax\|_Y}{\|x\|_X} = \sup_{x\in X, \|x\|_X \le 1} \|Ax\|_Y = \sup_{x\in X, \|x\|_X = 1} \|Ax\|_Y < \infty. \qquad \text{(A.46)}$$

The operator A is continuous in $x_0 \in X$ if for each $\varepsilon > 0$ there is a $\delta > 0$ such that for all $x \in X$ with $\|x - x_0\|_X < \delta$, it follows that $\|Ax - Ax_0\|_Y < \varepsilon$. The operator A is called a continuous operator if A is continuous for all $x \in X$. □

Remark A.61 (Equivalent Definition of a Continuous Operator) The operator $A : X \to Y$, X, Y Banach spaces, is continuous in $x_0 \in X$ if and only if for all sequences $\{x_n\}_{n=1}^{\infty}$, $x_n \in X$, with $x_n \to x_0$ it holds that $Ax_n \to Ax_0$ in Y. □

Lemma A.62 (Properties of Bounded Linear Operators, Kolmogorov and Fomīn (1975, §4.5.2, §4.5.3), Yosida (1995, p.43)) *Let X, Y be Banach spaces.*

(i) A bounded linear operator $A : X \to Y$ is continuous.
(ii) A continuous linear operator $A : X \to Y$ is bounded.
(iii) The set

$$\mathcal{L}(X,Y) = \{A : A \textit{ is a bounded linear operator from } X \textit{ to } Y\}$$

is a Banach space endowed with the norm (A.46).

Definition A.63 (Compact Operator) An operator $A : X \to Y$ is compact, if $A(x)$ is precompact in Y for every bounded set $\tilde{X} \subset X$. □

Theorem A.64 (Rank-Nullity Theorem) *Let $A : V \to W$ be a linear map between two finite-dimensional linear spaces V and W, then it holds*

$$\dim V = \dim(\ker(A)) + \dim(\text{range}(A)).$$

Definition A.65 (Linear Functional) A (real) linear functional f defined on a Banach space X is a linear operator with range$(f) \subset \mathbb{R}$. □

Definition A.66 (Bounded Bilinear Form, Coercive Bilinear Form, V-elliptic Bilinear Form) Let $b(\cdot,\cdot) : V \times V \to \mathbb{R}$ be a bilinear form on the Banach space V. Then, it is bounded if

$$|b(u,v)| \le M \|u\|_V \|v\|_V \quad \forall\, u, v \in V, M > 0, \tag{A.47}$$

where the constant M is independent of u and v. The bilinear form is coercive or V-elliptic if

$$b(u,u) \ge m \|u\|_V^2 \quad \forall\, u \in V, m > 0, \tag{A.48}$$

where the constant m is independent of u. □

Remark A.67 (Application to an Inner Product) Let V be a Hilbert space. Then the inner product $a(\cdot,\cdot)$ is a bounded and coercive bilinear form, since by the Cauchy–Schwarz inequality

$$|a(u,v)| \le \|u\|_V \|v\|_V \quad \forall\, u, v \in V,$$

and obviously $a(u, u) = \|u\|_V^2$. Hence, the constants can be chosen to be $M = 1$ and $m = 1$. □

Theorem A.68 (Banach's Fixed Point Theorem, Gilbarg and Trudinger (1983, Theorem 5.1)) *A contraction mapping T in a Banach space B has a unique fixed point, that is there exists a unique solution $x \in B$ of the equation $Tx = x$.*

The statement holds true if B is replaced by any closed subset, see Gilbarg and Trudinger (1983, p. 74).

Theorem A.69 (Brouwer's Fixed Point Theorem, Gilbarg and Trudinger (1983, Theorem 11.1)) *Let S be a compact convex set in a Banach space B and let T be a continuous mapping of S into itself. Then T has a fixed point, that is, $Tx = x$ for some $x \in S$.*

Theorem A.70 (Theorem of Banach on the Inverse Operator, Kolmogorov and Fomīn (1975, p. 225)) *Let $A : X \to Y$ be a bounded linear operator that defines a one-to-one mapping between the Banach spaces X and Y. Then, the inverse operator A^{-1} is bounded.*

Theorem A.71 (Closed Range Theorem of Banach, Yosida (1995, p. 205)) *Let X, Y be Banach spaces, let $A : X \to Y$ be a bounded linear operator, and let $A' : Y' \to X'$ be its dual. Then, the following statements are equivalent:*

(i) $\text{range}(A)$ *is closed in* Y,
(ii) $\text{range}(A) = \{y \in Y : \langle y', y\rangle_{Y',Y} = 0 \text{ for all } y' \in \ker(A')\}$,
(iii) $\text{range}(A')$ *is closed in* X,
(iv) $\text{range}(A') = \{x' \in X' : \langle x', x\rangle_{X',X} = 0 \text{ for all } x \in \ker(A)\}$.

Theorem A.72 (Hahn–Banach Theorem, Yosida (1995, Sect. IV.1), Triebel (1972, p. 67)) *Let X be a Banach space, let Y be a subspace of X, and let f be a bounded linear functional defined on Y. Then, there exists an extension g of f to X, where g is a linear functional with the same norm as f.*

Appendix B
Finite Element Methods

Remark B.1 (Contents) This appendix provides a short introduction into finite element methods. In particular, notations are introduced that are used throughout this monograph and finite element spaces are described that are of importance for the discretization of incompressible flow problems. □

B.1 The Ritz Method and the Galerkin Method

Remark B.2 (Contents) This section studies abstract problems in Hilbert spaces. The existence and uniqueness of solutions will be discussed. Approximating this solution with finite-dimensional spaces is called Ritz method or Galerkin method. Some basic properties of this method will be proved.

In this section, a Hilbert space V will be considered with inner product $a(\cdot,\cdot)$: $V \times V \to \mathbb{R}$ and norm $\|v\|_V = a(v,v)^{1/2}$. □

Theorem B.3 (Representation Theorem of Riesz) *Let $f \in V'$ be a continuous and linear functional, then there is a uniquely determined $u \in V$ with*

$$a(u,v) = f(v) \quad \forall\ v \in V. \tag{B.1}$$

In addition, u is the unique solution of the variational problem

$$F(v) = \frac{1}{2}a(v,v) - f(v) \to \min\ \forall\ v \in V. \tag{B.2}$$

Proof First, the existence of a solution u of the variational problem will be proved. Since f is continuous, it holds

$$|f(v)| \le c\,\|v\|_V \quad \forall\ v \in V,$$

V. John, *Finite Element Methods for Incompressible Flow Problems*, Springer Series in Computational Mathematics 51, DOI 10.1007/978-3-319-45750-5

from what follows that

$$F(v) \geq \frac{1}{2} \|v\|_V^2 - c \|v\|_V \geq -\frac{1}{2}c^2,$$

where in the second estimate the necessary criterion for a local minimum of the expression of the first bound is used. Hence, the function $F(\cdot)$ is bounded from below and

$$d = \inf_{v \in V} F(v)$$

exists.

Let $\{v_k\}_{k\in\mathbb{N}}$ be a sequence with $F(v_k) \to d$ for $k \to \infty$. A straightforward calculation (parallelogram identity in Hilbert spaces) gives

$$\|v_k - v_l\|_V^2 + \|v_k + v_l\|_V^2 = 2\|v_k\|_V^2 + 2\|v_l\|_V^2 .$$

Using the linearity of $f(\cdot)$ and $d \leq F(v)$ for all $v \in V$, one obtains

$$\begin{aligned}
&\|v_k - v_l\|_V^2 \\
&= 2\|v_k\|_V^2 + 2\|v_l\|_V^2 - 4\left\|\frac{v_k + v_l}{2}\right\|_V^2 - 4f(v_k) - 4f(v_l) + 8f\left(\frac{v_k + v_l}{2}\right) \\
&= 4F(v_k) + 4F(v_l) - 8F\left(\frac{v_k + v_l}{2}\right) \\
&\leq 4F(v_k) + 4F(v_l) - 8d \to 0
\end{aligned}$$

for $k, l \to \infty$. Hence $\{v_k\}_{k\in\mathbb{N}}$ is a Cauchy sequence. Because V is a complete space, there exists a limit u of this sequence with $u \in V$, see Definition A.5. Since $F(\cdot)$ is continuous, it is $F(u) = d$ and u is a solution of the variational problem.

In the next step, it will be shown that each solution of the variational problem (B.2) is also a solution of (B.1). It is

$$\begin{aligned}
\Phi(\varepsilon) = F(u + \varepsilon v) &= \frac{1}{2}a(u + \varepsilon v, u + \varepsilon v) - f(u + \varepsilon v) \\
&= \frac{1}{2}a(u, u) + \varepsilon a(u, v) + \frac{\varepsilon^2}{2}a(v, v) - f(u) - \varepsilon f(v).
\end{aligned}$$

If u is a minimum of the variational problem, then the function $\Phi(\varepsilon)$ has a local minimum at $\varepsilon = 0$. The necessary condition for a local minimum leads to

$$0 = \Phi'(0) = a(u, v) - f(v) \quad \text{for all } v \in V.$$

Finally, the uniqueness of the solution will be proved. It is sufficient to prove the uniqueness of the solution of Eq. (B.1). If the solution of (B.1) is unique, then the existence of two solutions of the variational problem (B.2) would be a contradiction to the fact proved in the previous step. Let u_1 and u_2 be two solutions of (B.1). Computing the difference of both equations gives

$$a(u_1 - u_2, v) = 0 \quad \text{for all } v \in V.$$

This equation holds, in particular, for $v = u_1 - u_2$. Hence, $\|u_1 - u_2\|_V = 0$, such that $u_1 = u_2$. ■

Theorem B.4 (Theorem of Lax–Milgram) *Let $b(\cdot,\cdot) : V \times V \to \mathbb{R}$ be a bounded and coercive bilinear form on the Hilbert space V. Then, for each bounded linear functional $f \in V'$ there is exactly one $u \in V$ with*

$$b(u, v) = f(v) \quad \forall\ v \in V. \tag{B.3}$$

Proof One defines linear operators $T, T' : V \to V$ by

$$a(Tu, v) = b(u, v)\ \forall\ v \in V, \quad a(T'u, v) = b(v, u)\ \forall\ v \in V. \tag{B.4}$$

Since $b(u,\cdot)$ and $b(\cdot,u)$ are continuous linear functionals on V, it follows from Theorem B.3 that the elements Tu and $T'u$ exist and they are defined uniquely. The operators satisfy the relation

$$a(Tu, v) = b(u, v) = a(T'v, u) = a(u, T'v), \tag{B.5}$$

T' is called adjoint operator of T. Setting $v = Tu$ in (B.4) and using the boundedness of $b(\cdot,\cdot)$ yields

$$\|Tu\|_V^2 = a(Tu, Tu) = b(u, Tu) \le M \|u\|_V \|Tu\|_V \implies \|Tu\|_V \le M \|u\|_V$$

for all $u \in V$. Hence, T is bounded. Since T is linear, it follows that T is continuous, see Lemma A.62. Using the same argument, one shows that T' is also bounded and continuous.

Define the bilinear form

$$d(u, v) := a(TT'u, v) = a(T'u, T'v) \quad \forall\ u, v \in V,$$

where (B.5) was used. Hence, this bilinear form is symmetric. Using the coercivity of $b(\cdot,\cdot)$ and the Cauchy–Schwarz inequality (A.10) gives

$$\begin{aligned} m^2 \|v\|_V^4 \le b(v, v)^2 = a(T'v, v)^2 \le \|v\|_V^2 \left\|T'v\right\|_V^2 = \|v\|_V^2\, a(T'v, T'v) \\ = \|v\|_V^2\, d(v, v). \end{aligned}$$

Applying now the boundedness of $a(\cdot,\cdot)$ and of T' yields

$$m^2 \|v\|_V^2 \leq d(v,v) = a(T'v, T'v) = \|T'v\|_V^2 \leq M \|v\|_V^2 . \tag{B.6}$$

Hence, $d(\cdot,\cdot)$ is also coercive and, since it is symmetric, it defines an inner product on V. From (B.6), one has that the norm induced by $d(v,v)^{1/2}$ is equivalent to the norm $\|v\|_V$. From Theorem B.3, it follows that there is a exactly one $w \in V$ with

$$d(w,v) = f(v) \quad \forall\ v \in V.$$

Inserting $u = T'w$ in (B.3) gives with (B.4)

$$b(T'w,v) = a(TT'w,v) = d(w,v) = f(v) \quad \forall\ v \in V$$

and consequently, $u = T'w$ is a solution of (B.3).

The uniqueness of the solution is proved analogously as in the symmetric case. ■

Remark B.5 (Basic Idea of the Ritz Method) For approximating the solution of (B.2) or (B.1) with a numerical method, it will be assumed that V has a countable orthonormal basis (Schauder basis), i.e., V is a separable Hilbert space. Then, using Parseval's equality, one finds that there are finite-dimensional subspaces $V_1, V_2, \ldots \subset V$ with $\dim V_k = k$, which have the following property: for each $u \in V$ and each $\varepsilon > 0$ there is a $K \in \mathbb{N}$ and a $u_k \in V_k$ with

$$\|u - u_k\|_V \leq \varepsilon \quad \forall\ k \geq K. \tag{B.7}$$

Note that it is not required that there holds an inclusion of the form $V_k \subset V_{k+1}$.

The Ritz approximation of (B.2) and (B.1) is defined by: Find $u_k \in V_k$ with

$$a(u_k, v_k) = f(v_k) \quad \forall\ v_k \in V_k. \tag{B.8}$$

□

Lemma B.6 (Existence and Uniqueness of a Solution of (B.8)) *There exists exactly one solution of (B.8).*

Proof Finite-dimensional subspaces of Hilbert spaces are Hilbert spaces as well. For this reason, one can apply the representation theorem of Riesz, Theorem B.3, to (B.8) which gives the statement of the lemma. In addition, the solution of (B.8) solves a minimization problem on V_k. ■

Lemma B.7 (Best Approximation Property) *The solution of (B.8) is the best approximation of u in V_k, i.e., it is*

$$\|u - u_k\|_V = \inf_{v_k \in V_k} \|u - v_k\|_V . \tag{B.9}$$

Proof Since $V_k \subset V$, one can use the test functions from V_k in the weak equation (B.1). Then, the difference of (B.1) and (B.8) gives the orthogonality, the so-called Galerkin orthogonality,

$$a(u - u_k, v_k) = 0 \quad \forall\ v_k \in V_k. \tag{B.10}$$

Hence, the error $u - u_k$ is orthogonal to the space V_k: $u - u_k \perp V_k$. That means, u_k is the orthogonal projection of u onto V_k with respect of the inner product of V.

Let now $w_k \in V_k$ be an arbitrary element, then it follows with the Galerkin orthogonality (B.10) and the Cauchy–Schwarz inequality (A.10) that

$$\begin{aligned} \|u - u_k\|_V^2 &= a(u - u_k, u - u_k) = a(u - u_k, u - \underbrace{(u_k - w_k)}_{v_k}) = a(u - u_k, u - v_k) \\ &\le \|u - u_k\|_V \, \|u - v_k\|_V \, . \end{aligned}$$

Since $w_k \in V_k$ was arbitrary, also $v_k \in V_k$ is arbitrary. If $\|u - u_k\|_V > 0$, division by $\|u - u_k\|_V$ gives the statement of the lemma. If $\|u - u_k\|_V = 0$, the statement of the lemma is trivially true. ■

Theorem B.8 (Convergence of the Ritz Approximation) *The Ritz approximation converges*

$$\lim_{k \to \infty} \|u - u_k\|_V = 0.$$

Proof The best approximation property (B.9) and property (B.7) give

$$\|u - u_k\|_V = \inf_{v_k \in V_k} \|u - v_k\|_V \le \varepsilon$$

for each $\varepsilon > 0$ and $k \ge K(\varepsilon)$. Hence, the convergence is proved. ■

Remark B.9 (Formulation of the Ritz Method as Linear System of Equations) One can use an arbitrary basis $\{\phi_i\}_{i=1}^k$ of V_k for the computation of u_k. First of all, the equation for the Ritz approximation (B.8) is satisfied for all $v_k \in V_k$ if and only if it is satisfied for each basis function ϕ_i. This statement follows from the linearity of both sides of the equation with respect to the test function and from the fact that each function $v_k \in V_k$ can be represented as linear combination of the basis functions. Let $v_k = \sum_{i=i}^k \alpha_i \phi_i$, then from (B.8), it follows that

$$a(u_k, v_k) = \sum_{k=1}^{k} \alpha_i a(u_k, \phi_i) = \sum_{k=1}^{k} \alpha_i f(\phi_i) = f(v_k).$$

This equation is satisfied if $a(u_k, \phi_i) = f(\phi_i)$, $i = 1, \ldots, k$. On the other hand, if (B.8) hold, then it holds in particular for each basis function ϕ_i.

One uses as ansatz for the solution also a linear combination of the basis functions

$$u_k = \sum_{j=1}^{k} u^j \phi_j$$

with unknown coefficients $u^j \in \mathbb{R}$. Using as test functions now the basis functions yields

$$\sum_{j=1}^{k} a(u^j \phi_j, \phi_i) = \sum_{j=1}^{k} a(\phi_j, \phi_i) u^j = f(\phi_i), \quad i = 1, \ldots, k.$$

This equation is equivalent to the linear system of equations $A\underline{u} = \underline{f}$, where

$$A = \left(a_{ij}\right)_{i,j=1}^{k} = a(\phi_j, \phi_i)_{i,j=1}^{k}$$

is called stiffness matrix. Note that the order of the indices is different for the entries of the matrix and the arguments of the inner product. The right-hand side is a vector of length k with the entries $f_i = f(\phi_i)$, $i = 1, \ldots, k$.

Using the one-to-one mapping between the coefficient vector $(v^1, \ldots, v^k)^T$ and the element $v_k = \sum_{i=1}^{k} v^i \phi_i$, one can show that the matrix A is symmetric and positive definite

$$\begin{aligned} A = A^T \quad &\Longleftrightarrow \quad a(v, w) = a(w, v) \quad \forall\, v, w \in V_k, \\ \underline{x}^T A \underline{x} > 0 \text{ for } \underline{x} \neq \underline{0} \quad &\Longleftrightarrow \quad a(v, v) > 0 \quad \forall\, v \in V_k, v \neq 0. \end{aligned}$$

□

Remark B.10 (The Case of a Bounded and Coercive Bilinear Form) If $b(\cdot, \cdot)$ is bounded and coercive, but not symmetric, it is possible to approximate the solution of (B.3) with the same idea as for the Ritz method. In this case, it is called Galerkin method. The discrete problem consists in finding $u_k \in V_k$ such that

$$b(u_k, v_k) = f(v_k) \quad \forall\, v_k \in V_k. \tag{B.11}$$

□

Lemma B.11 (Existence and Uniqueness of a Solution of (B.11)) *There is exactly one solution of (B.11).*

Proof The statement of the lemma follows directly from the Theorem of Lax–Milgram, Theorem B.4. ■

Lemma B.12 (Lemma of Cea, Error Estimate) *Let $b : V \times V \to \mathbb{R}$ be a bounded and coercive bilinear form on the Hilbert space V and let $f \in V'$ be a bounded linear* *functional. Let u be the solution of* (B.3) *and let u_k be the solution of* (B.11)*, then the following error estimate holds*

$$\|u - u_k\|_V \leq \frac{M}{m} \inf_{v_k \in V_k} \|u - v_k\|_V\,, \tag{B.12}$$

where the constants M and m are given in (A.47) *and* (A.48)*.*

Proof Considering the difference of the continuous equation (B.3) and the discrete equation (B.11), one obtains the error equation

$$b(u - u_k, v_k) = 0 \quad \forall\ v_k \in V_k,$$

which is also called Galerkin orthogonality. With (A.48), the Galerkin orthogonality, and (A.47), it follows that

$$\begin{aligned}\|u - u_k\|_V^2 &\leq \frac{1}{m} b(u - u_k, u - u_k) = \frac{1}{m} b(u - u_k, u - v_k)\\ &\leq \frac{M}{m}\, \|u - u_k\|_V\, \|u - v_k\|_V \quad \forall\ v_k \in V_k,\end{aligned}$$

from what the statement of the lemma follows immediately. ■

Remark B.13 (On the Best Approximation Error) It follows from estimate (B.12) that the error is bounded by a multiple of the best approximation error, where the factor depends on properties of the bilinear form $b(\cdot,\cdot)$. Thus, concerning error estimates for concrete finite-dimensional spaces, the study of the best approximation error will be of importance. □

Remark B.14 (The Corresponding Linear System of Equations) The corresponding linear system of equations is derived analogously to the symmetric case. The system matrix is still positive definite but not symmetric. □

Lemma B.15 (Inf-Sup Criterion for Finite-Dimensional Spaces) *Let V^h be a finite-dimensional space with inner product $(\cdot,\cdot)_{V^h}$ and induced norm $\|\cdot\|_{V^h} = (\cdot,\cdot)_{V^h}^{1/2}$. Consider a bilinear form $a : V^h \times V^h \to \mathbb{R}$ and a linear functional $f : V^h \to \mathbb{R}$. Then, the problem to find $u^h \in V^h$ such that*

$$a\left(u^h, v^h\right) = f\left(v^h\right) \tag{B.13}$$

has a unique solution for all $f(\cdot)$ if and only if

$$\inf_{w^h \in V^h \setminus \{0\}} \sup_{v^h \in V^h \setminus \{0\}} \frac{a\left(v^h, w^h\right)}{\|v^h\|_{V^h}\, \|w^h\|_{V^h}} \geq \beta_{\text{is}}^h > 0. \tag{B.14}$$

Proof Denote by $n \in \mathbb{N}$ the dimension of V^h and let $\{\varphi_i^h\}_{i=1}^n$ be a basis of V^h. Then there are representations

$$v^h = \sum_{i=1}^n v_i \varphi_i^h, \quad w^h = \sum_{i=1}^n w_i \varphi_i^h,$$

with $\underline{v} = (v_1, \ldots, v_n)^T$, $\underline{w} = (w_1, \ldots, w_n)^T$ and there are matrices $A, M \in \mathbb{R}^{n \times n}$ such that

$$a\left(v^h, w^h\right) = \underline{v}^T A^T \underline{w}, \quad \left\|v^h\right\|_{V^h} = \left(\underline{v}^T M^T \underline{v}\right)^{1/2}, \quad \left\|w^h\right\|_{V^h} = \left(\underline{w}^T M^T \underline{w}\right)^{1/2}.$$

The matrix M is symmetric and positive definite. The inf-sup condition (B.14) can be written in the form

$$\inf_{\underline{w} \in \mathbb{R}^n \setminus \{\underline{0}\}} \sup_{\underline{v} \in \mathbb{R}^n \setminus \{\underline{0}\}} \frac{\underline{v}^T A^T \underline{w}}{\left(\underline{v}^T M \underline{v}\right)^{1/2} \left(\underline{w}^T M \underline{w}\right)^{1/2}} \geq \beta_{\text{is}}^h > 0. \tag{B.15}$$

Problem (B.13) has a unique solution for all f if and only if the matrix A is non-singular.

(B.15) *Holds* $\Longrightarrow$ *A is Non-singular* Assume that (B.15) holds and A is a singular matrix. Then there is a vector $\underline{v} \neq \underline{0}$ such that $A\underline{v} = \underline{0}$ or equivalently $\underline{v}^T A^T = \underline{0}^T$. Hence, the supremum in (B.15) is zero such that (B.15) cannot hold, which is a contradiction to the assumption.

A is Non-singular $\Longrightarrow$ (B.15) *Holds* With A also A^T is non-singular. Then, for each $\underline{v} \in \mathbb{R}^n \setminus \{\underline{0}\}$ there is a unique $\underline{w} \in \mathbb{R}^n \setminus \{\underline{0}\}$ such that $\underline{v} = M^{-1} A^T \underline{w}$, since M and A^T are non-singular matrices. Inserting this vector in (B.15) and using the symmetry and positive definiteness of M gives

$$\begin{aligned}
&\inf_{\underline{w} \in \mathbb{R}^n \setminus \{\underline{0}\}} \frac{\underline{w}^T A M^{-1} A^T \underline{w}}{\left(\underline{w}^T A M^{-1} M M^{-1} A^T \underline{w}\right)^{1/2} \left(\underline{w}^T M \underline{w}\right)^{1/2}} \\
&= \inf_{\underline{w} \in \mathbb{R}^n \setminus \{\underline{0}\}} \frac{\underline{w}^T A M^{-1} A^T \underline{w}}{\left(\underline{w}^T A M^{-1} A^T \underline{w}\right)^{1/2} \left(\underline{w}^T M \underline{w}\right)^{1/2}} \\
&= \inf_{\underline{w} \in \mathbb{R}^n \setminus \{\underline{0}\}} \left(\frac{\underline{w}^T A M^{-1} A^T \underline{w}}{\underline{w}^T M \underline{w}}\right)^{1/2} \\
&= \inf_{\underline{w} \in \mathbb{R}^n \setminus \{\underline{0}\}} \left(\frac{\left(\underline{w} M^{1/2}\right)^T \left(M^{-1/2} A M^{-1/2}\right) \left(M^{-T/2} A^T M^{-T/2}\right) \left(M^{1/2} \underline{w}\right)}{\left(\underline{w} M^{1/2}\right)^T \left(M^{1/2} \underline{w}\right)}\right)^{1/2}.
\end{aligned}$$

Hence, one obtains a Rayleigh quotient. From Lemma A.19, it is known that the infimum of a Rayleigh quotient is attained and it is the smallest eigenvalue of the eigenvalue problem

$$\left(M^{-1/2}AM^{-1/2}\right)\left(M^{-T/2}A^TM^{-T/2}\right)\left(M^{1/2}\underline{w}\right) = \lambda\left(M^{1/2}\underline{w}\right).$$

This problem is an eigenvalue problem for a symmetric matrix, hence all eigenvalues are real. Since

$$\begin{aligned}
&\underline{w}^T\left(M^{-1/2}AM^{-1/2}\right)\left(M^{-T/2}A^TM^{-T/2}\right)\underline{w} \\
&= \left(\left(M^{-T/2}A^TM^{-T/2}\right)\underline{w}\right)^T\left(\left(M^{-T/2}A^TM^{-T/2}\right)\underline{w}\right) \\
&= \left\|\left(M^{-T/2}A^TM^{-T/2}\right)\underline{w}\right\|_2^2 \geq 0,
\end{aligned}$$

the matrix is positive semi-definite, such that all eigenvalues are non-negative. Finally, since A and M are non-singular matrices, the matrix of this eigenvalue problem is non-singular and all eigenvalues are positive. Hence, there is a positive constant β_{is}^h such that

$$\inf_{\underline{w}\in\mathbb{R}^n\setminus\{\underline{0}\}} \frac{\underline{v}^TA^T\underline{w}}{\left(\underline{v}^TM\underline{v}\right)^{1/2}\left(\underline{w}^TM\underline{w}\right)^{1/2}} \geq \beta_{\mathrm{is}}^h$$

with $\underline{v} = M^{-1}A^T\underline{w}$. Taking now the supremum with respect to $\underline{v}$ might only increase the left-hand side and (B.15) follows. ■

B.2 Finite Element Spaces

Remark B.16 (Mesh Cells, Faces, Edges, Vertices) A mesh cell K is a compact polyhedron in $\mathbb{R}^d$, $d \in \{2, 3\}$, whose interior is not empty. The boundary ∂K of K consists of m-dimensional linear manifolds (points, pieces of straight lines, pieces of planes), $0 \leq m \leq d-1$, which are called m-faces. The 0-faces are the vertices of the mesh cell, the 1-faces are the edges, and the $(d-1)$-faces are just called faces. □

Remark B.17 (Finite-Dimensional Spaces Defined on K) Let $s \in \mathbb{N}$. Finite element methods use finite-dimensional spaces $P(K) \subset C^s(K)$ that are defined on K. In general, $P(K)$ consists of polynomials. The dimension of $P(K)$ will be denoted by $\dim P(K) = N_K$. □

Remark B.18 (Linear Functionals Defined on $P(K)$, Nodal Functionals) For the definition of finite elements, linear functional that are defined on $P(K)$ are of importance. These functionals are called nodal functionals.

Consider linear and continuous functionals $\Phi_{K,1}, \ldots, \Phi_{K,N_K} : C^s(K) \to \mathbb{R}$ which are linearly independent. There are different types of functionals that can be utilized in finite element methods:

- point values: $\Phi(v) = v(\boldsymbol{x}), \boldsymbol{x} \in K$,
- point values of a first partial derivative: $\Phi(v) = \partial_i v(\boldsymbol{x}), \boldsymbol{x} \in K$,
- point values of the normal derivative on a face E of K: $\Phi(v) = \nabla v(\boldsymbol{x}) \cdot \boldsymbol{n}_E$, $\boldsymbol{n}_E$ is the outward pointing unit normal vector on E,
- integral mean values on K: $\Phi(v) = \frac{1}{|K|} \int_K v(\boldsymbol{x})\, d\boldsymbol{x}$,
- integral mean values on faces E: $\Phi(v) = \frac{1}{|E|} \int_E v(\boldsymbol{s})\, d\boldsymbol{s}$.

The smoothness parameter s has to be chosen in such a way that the functionals $\Phi_{K,1}, \ldots, \Phi_{K,N_K}$ are continuous. If, e.g., a functional requires the evaluation of a partial derivative or a normal derivative, then one has to choose at least $s = 1$. For the other functionals given above, $s = 0$ is sufficient. □

Definition B.19 (Unisolvence of $P(K)$ with Respect to the Functionals $\Phi_{K,1}, \ldots, \Phi_{K,N_K}$) The space $P(K)$ is called unisolvent with respect to the functionals $\Phi_{K,1}, \ldots, \Phi_{K,N_K}$ if there is for each $\underline{a} \in \mathbb{R}^{N_K}$, $\underline{a} = (a_1, \ldots, a_{N_K})^T$, exactly one $p \in P(K)$ with

$$\Phi_{K,i}(p) = a_i, \quad 1 \le i \le N_K.$$

□

Remark B.20 (Local Basis) Unisolvence means that for each vector $\underline{a} \in \mathbb{R}^{N_K}$, $\underline{a} = (a_1, \ldots, a_{N_K})^T$, there is exactly one element in $P(K)$ such that a_i is the image of the ith functional, $i = 1, \ldots, N_K$.

Choosing in particular the Cartesian unit vectors for $\underline{a}$, then it follows from the unisolvence that a set $\{\phi_{K,i}\}_{i=1}^{N_K}$ exists with $\phi_{K,i} \in P(K)$ and

$$\Phi_{K,i}(\phi_{K,j}) = \delta_{ij}, \quad i,j = 1, \ldots, N_K.$$

Consequently, the set $\{\phi_{K,i}\}_{i=1}^{N_K}$ forms a basis of $P(K)$. This basis is called local basis. □

Remark B.21 (Transform of an Arbitrary Basis to the Local Basis) If an arbitrary basis $\{p_i\}_{i=1}^{N_K}$ of $P(K)$ is known, then the local basis can be computed by solving a linear system of equations. To this end, represent the local basis in terms of the known basis

$$\phi_{K,j} = \sum_{k=1}^{N_K} c_{jk} p_k, \quad c_{jk} \in \mathbb{R},\ j = 1, \ldots, N_K,$$

with unknown coefficients c_{jk}. Applying the definition of the local basis leads to the linear system of equations

$$\Phi_{K,i}(\phi_{K,j}) = \sum_{k=1}^{N_K} c_{jk} a_{ik} = \delta_{ij}, \quad i,j = 1,\ldots,N_K, \quad a_{ik} = \Phi_{K,i}(p_k).$$

Because of the unisolvence, the matrix $A = (a_{ij})$ is non-singular and the coefficients c_{jk} are determined uniquely. □

Remark B.22 (Triangulation, Grid, Mesh, Grid Cell) For the definition of global finite element spaces, a decomposition of the domain Ω into polyhedra K is needed. This decomposition is called triangulation $\mathcal{T}^h$ and the polyhedra K are called mesh cells. The union of the polyhedra is called grid or mesh.

A triangulation is called admissible, see the definition in Ciarlet (1978, p. 38, p. 51), if:

- It holds $\overline{\Omega} = \cup_{K\in\mathcal{T}^h} K$.
- Each mesh cell $K \in \mathcal{T}^h$ is closed and the interior $\mathring{K}$ is non-empty.
- For distinct mesh cells K_1 and K_2 there holds $\mathring{K}_1 \cap \mathring{K}_2 = \emptyset$.
- For each $K \in \mathcal{T}^h$, the boundary ∂K is Lipschitz continuous.
- The intersection of two mesh cells is either empty or a common m-face, $m \in \{0,\ldots,d-1\}$.

□

Remark B.23 (Global and Local Functionals) Let $\Phi_1,\ldots,\Phi_N : C^s(\overline{\Omega}) \to \mathbb{R}$ continuous linear functionals of the same types as given in Remark B.18. The restriction of the functionals to $C^s(K)$ defines a set of local functionals $\Phi_{K,1},\ldots,\Phi_{K,N_K}$, where it is assumed that the local functionals are unisolvent on $P(K)$. The union of all mesh cells K_j, for which there is a $p \in P(K_j)$ with $\Phi_i(p) \neq 0$, will be denoted by ω_i. □

Example B.24 (On Subdomains ω_i) Consider the two-dimensional case and let Φ_i be defined as nodal value of a function in $\boldsymbol{x} \in K$. If $\boldsymbol{x} \in \mathring{K}$, then $\omega_i = K$. In the case that $\boldsymbol{x}$ is on a face of K but not in a vertex, then ω_i is the union of K and the other mesh cell whose boundary contains this face. Last, if $\boldsymbol{x}$ is a vertex of K, then ω_i is the union of all mesh cells that possess this vertex. □

Definition B.25 (Finite Element Space, Global Basis) A function $v(\boldsymbol{x})$ defined on Ω with $v|_K \in P(K)$ for all $K \in \mathcal{T}^h$ is called continuous with respect to the functional $\Phi_i : \Omega \to \mathbb{R}$ if

$$\Phi_i(v|_{K_1}) = \Phi_i(v|_{K_2}), \quad \forall\, K_1, K_2 \in \omega_i.$$

The space

$$S = \Big\{ v \in L^\infty(\Omega) \ : \ v|_K \in P(K) \text{ and } v \text{ is continuous with respect to}$$
$$\Phi_i, i = 1, \ldots, N \Big\}$$

is called finite element space.

The global basis $\{\phi_j\}_{j=1}^N$ of S is defined by the condition

$$\phi_j \in S, \quad \Phi_i(\phi_j) = \delta_{ij}, \quad i,j = 1, \ldots, N.$$

□

Remark B.26 (On Global Basis Functions) A global basis function coincides on each mesh cell with a local basis function. This property implies the uniqueness of the global basis functions.

Whether the continuity with respect to $\{\Phi_i\}_{i=1}^N$ implies the continuity of the finite element functions depends on the functionals that define the finite element space.

□

Definition B.27 (Parametric Finite Elements) Let $\hat{K}$ be a reference mesh cell with the local space $P(\hat{K})$, the local functionals $\hat{\Phi}_1, \ldots, \hat{\Phi}_{\hat{N}}$, and a class of bijective mappings $\{F_K \ : \ \hat{K} \to K\}$. A finite element space is called a parametric finite element space if:

- The images $\{K\}$ of $\{F_K\}$ form the set of mesh cells.
- The local spaces are given by

$$P(K) = \Big\{ p \ : \ p = \hat{p} \circ F_K^{-1}, \hat{p} \in \hat{P}(\hat{K}) \Big\}. \tag{B.16}$$

- The local functionals are defined by

$$\Phi_{K,i}(v(\boldsymbol{x})) = \hat{\Phi}_i\left(v(F_K(\hat{\boldsymbol{x}}))\right), \tag{B.17}$$

  where $\hat{\boldsymbol{x}} = (\hat{x}_1, \ldots, \hat{x}_d)^T$ are the coordinates of the reference mesh cell and it holds $\boldsymbol{x} = F_K(\hat{\boldsymbol{x}})$.

□

Remark B.28 (Motivations for Using Parametric Finite Elements) Definition B.25 of finite elements spaces is very general. For instance, different types of mesh cells are allowed. However, as well the finite element theory as the implementation of finite element methods become much simpler if only parametric finite elements are considered. □

B.3 Finite Elements on Simplices

Definition B.29 (d-Simplex) A d-simplex $K \subset \mathbb{R}^d$ is the convex hull of $(d+1)$ points $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{d+1} \in \mathbb{R}^d$ which form the vertices of K. □

Remark B.30 (On d-Simplices) It will be always assumed that the simplex is not degenerated, i.e., its d-dimensional measure is positive. This property is equivalent to the non-singularity of the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1,d+1} \\ a_{21} & a_{22} & \ldots & a_{2,d+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d1} & a_{d2} & \ldots & a_{d,d+1} \\ 1 & 1 & \ldots & 1 \end{pmatrix},$$

where $\boldsymbol{a}_i = (a_{1i}, a_{2i}, \ldots, a_{di})^T$, $i = 1, \ldots, d+1$.

For $d = 2$, the simplices are the triangles and for $d = 3$ they are the tetrahedra. □

Definition B.31 (Barycentric Coordinates) Since K is the convex hull of the points $\{\boldsymbol{a}_i\}_{i=1}^{d+1}$, the parametrization of K with a convex combination of the vertices reads as follows

$$K = \left\{ \boldsymbol{x} \in \mathbb{R}^d \ : \ \boldsymbol{x} = \sum_{i=1}^{d+1} \lambda_i \boldsymbol{a}_i, \ 0 \le \lambda_i \le 1, \ \sum_{i=1}^{d+1} \lambda_i = 1 \right\}.$$

The coefficients $\lambda_1, \ldots, \lambda_{d+1}$ are called barycentric coordinates of $\boldsymbol{x} \in K$. □

Remark B.32 (On Barycentric Coordinates) From the definition, it follows that the barycentric coordinates are the solution of the linear system of equations

$$\sum_{i=1}^{d+1} a_{ji} \lambda_i = x_j, \quad 1 \le j \le d, \quad \sum_{i=1}^{d+1} \lambda_i = 1.$$

Since the system matrix is non-singular, see Remark B.30, the barycentric coordinates are determined uniquely.

The barycentric coordinates of the vertex $\boldsymbol{a}_i$, $i = 1, \ldots, d+1$, of the simplex are $\lambda_i = 1$ and $\lambda_j = 0$ if $i \neq j$. Since $\lambda_i(\boldsymbol{a}_j) = \delta_{ij}$, the barycentric coordinate λ_i can be identified with the linear function that has the value 1 in the vertex $\boldsymbol{a}_i$ and that vanishes in all other vertices $\boldsymbol{a}_j$ with $j \neq i$.

The barycenter of the simplex is given by

$$S_K = \frac{1}{d+1}\sum_{i=1}^{d+1} \boldsymbol{a}_i = \sum_{i=1}^{d+1} \frac{1}{d+1}\boldsymbol{a}_i.$$

Hence, its barycentric coordinates are $\lambda_i = 1/(d+1)$, $i = 1, \ldots, d+1$. □

Remark B.33 (Simplicial Reference Mesh Cells) A commonly used reference mesh cell for triangles and tetrahedra is the unit simplex

$$\hat{K} = \left\{ \hat{\boldsymbol{x}} \in \mathbb{R}^d \ : \ \sum_{i=1}^{d} \hat{x}_i \le 1,\ \hat{x}_i \ge 0,\ i = 1, \ldots, d \right\},$$

see Fig. B.1. The class $\{F_K\}$ of admissible mappings are the bijective affine mappings

$$F_K\hat{\boldsymbol{x}} = B_K\hat{\boldsymbol{x}} + \boldsymbol{b}, \quad B_K \in \mathbb{R}^{d\times d},\ \det(B_K) \neq 0,\ \boldsymbol{b} \in \mathbb{R}^d. \tag{B.18}$$

The images of these mappings generate the set of the non-degenerated simplices $\{K\} \subset \mathbb{R}^d$. □

Definition B.34 (Affine Family of Simplicial Finite Elements) Given a simplicial reference mesh cell $\hat{K}$, affine mappings $\{F_K\}$, and an unisolvent set of functionals on $\hat{K}$. Using (B.16) and (B.17), one obtains a local finite element space on each non-degenerated simplex. The set of these local spaces is called affine family of simplicial finite elements. □

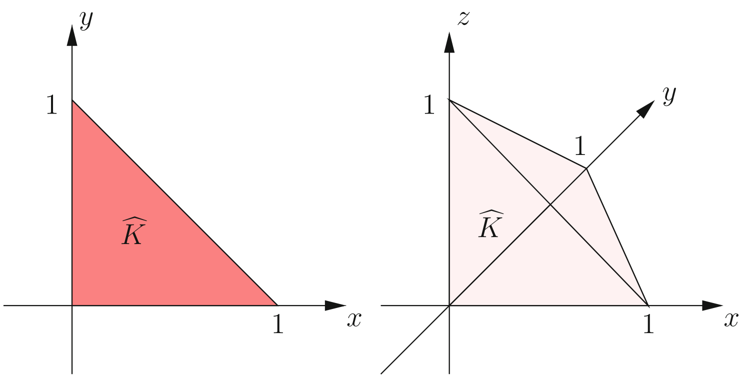

Fig. B.1 The unit simplices in two and three dimensions

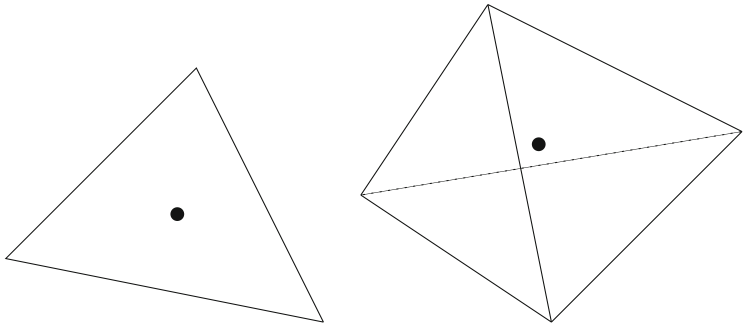

Fig. B.2 The finite element $P_0(K)$

Definition B.35 (Polynomial Space P_k) Let $\boldsymbol{x} = (x_1, \ldots, x_d)^T$, $k \in \mathbb{N} \cup \{0\}$, and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)^T$. Then, the polynomial space P_k is given by

$$P_k = \text{span} \left\{ \prod_{i=1}^{d} x_i^{\alpha_i} = \boldsymbol{x}^{\boldsymbol{\alpha}} \ : \ \alpha_i \in \mathbb{N} \cup \{0\} \ \text{ for } \ i = 1, \ldots, d, \ \sum_{i=1}^{d} \alpha_i \le k \right\} .$$

□

Remark B.36 (Lagrangian Finite Elements) In many examples given below, the linear functionals on the reference mesh cell $\hat{K}$ are the values of the polynomials with the same barycentric coordinates as on the general mesh cell K. Finite elements whose linear functionals are values of the polynomials on certain points in K are called Lagrangian finite elements. □

Example B.37 (P_0: Piecewise Constant Finite Element) The piecewise constant finite element space consists of discontinuous functions. The linear functional is the value of the polynomial in the barycenter of the mesh cell, see Fig. B.2. It is $\dim P_0(K) = 1$. □

Example B.38 (P_1: Conforming Piecewise Linear Finite Element) This finite element space is a subspace of $C(\overline{\Omega})$. The linear functionals are the values of the function in the vertices of the mesh cells, see Fig. B.3. It follows that $\dim P_1(K) = d + 1$.

The local basis for the functionals $\{\Phi_i(v) = v(\boldsymbol{a}_i), i = 1, \ldots, d + 1\}$, is $\{\lambda_i\}_{i=1}^{d+1}$ since $\Phi_i(\lambda_j) = \delta_{ij}$, compare Remark B.32. Since a local basis exists, the functionals are unisolvent with respect to the polynomial space $P_1(K)$.

Now, it will be shown that the corresponding finite element space consists of continuous functions. Let K_1, K_2 be two mesh cells with the common face E and let $v \in P_1(= S)$. The restriction of v_{K_1} on E is a linear function on E as well as the restriction of v_{K_2} on E. It has to be shown that both linear functions are identical. A linear function on the $(d-1)$-dimensional face E is uniquely determined with

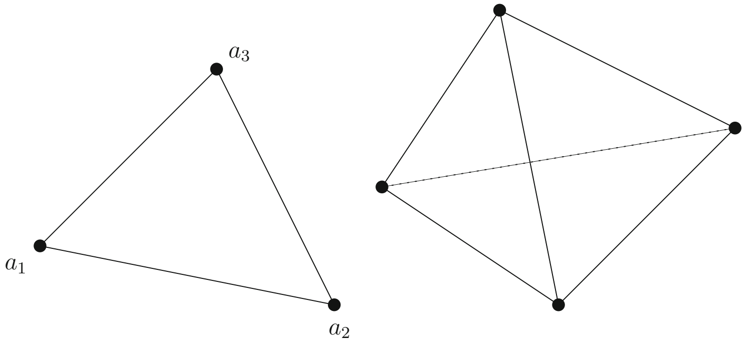

Fig. B.3 The finite element $P_1(K)$

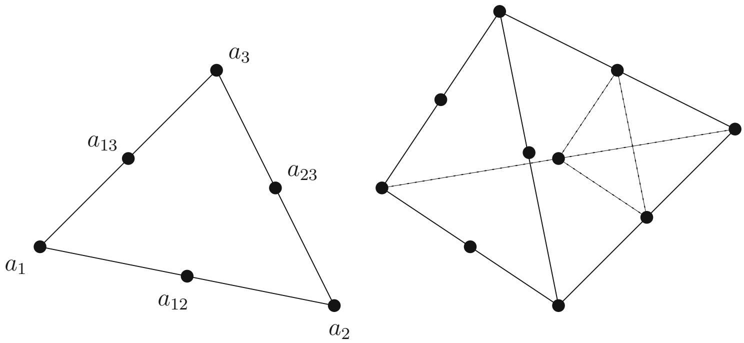

Fig. B.4 The finite element $P_2(K)$

d linearly independent functionals that are defined on E. These functionals can be chosen to be the values of the function in the d vertices of E. The functionals in S are continuous by the definition of S. Thus, it must hold that both restrictions on E have the same values in the vertices of E. Hence, it is $v_{K_1}|_E = v_{K_2}|_E$ and the functions from P_1 are continuous. □

Example B.39 (P_2: Conforming Piecewise Quadratic Finite Element) This finite element space is also a subspace of $C(\overline{\Omega})$. It consists of piecewise quadratic functions. The functionals are the values of the functions in the $d+1$ vertices of the mesh cell and the values of the functions in the centers of the edges, see Fig. B.4. Since each vertex is connected to each other vertex, there are $\sum_{i=1}^{d} i = d(d+1)/2$ edges. Hence, it follows that $\dim P_2(K) = (d+1)(d+2)/2$.

The part of the local basis that belongs to the functionals $\{\Phi_i(v) = v(\boldsymbol{a}_i),\ i = 1, \ldots, d+1\}$, is given by

$$\{\phi_i(\lambda) = \lambda_i(2\lambda_i - 1), \quad i = 1, \ldots, d+1\}.$$

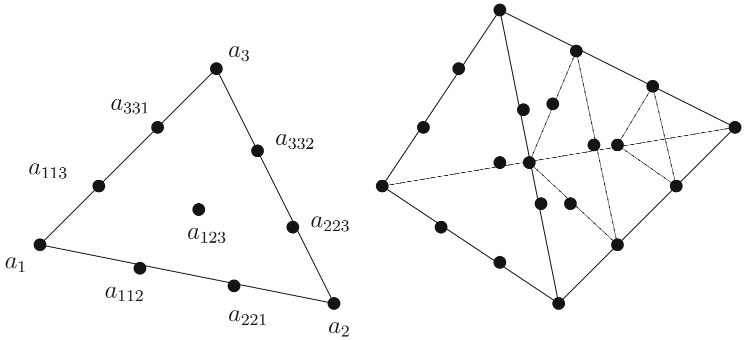

Fig. B.5 The finite element $P_3(K)$

Denote the center of the edges between the vertices $\boldsymbol{a}_i$ and $\boldsymbol{a}_j$ by $\boldsymbol{a}_{ij}$. The corresponding part of the local basis is given by

$$\{\phi_{ij} = 4\lambda_i\lambda_j, \quad i,j = 1,\ldots,d+1,\ i<j\}.$$

The unisolvence follows from the fact that there exists a local basis. The continuity of the corresponding finite element space is shown in the same way as for the P_1 finite element. The restriction of a quadratic function defined in a mesh cell to a face E is a quadratic function on that face. Hence, the function on E is determined uniquely with $d(d+1)/2$ linearly independent functionals on E.

The functions ϕ_{ij} are called in two dimensions edge bubble functions. □

Example B.40 (P_3: Conforming Piecewise Cubic Finite Element) This finite element space consists of continuous piecewise cubic functions. It is a subspace of $C(\overline{\Omega})$. The functionals in a mesh cell K are defined to be the values in the vertices ($(d+1)$ values), two values on each edge (dividing the edge in three parts of equal length) ($2\sum_{i=1}^{d} i = d(d+1)$ values), and the values in the barycenter of the 2-faces of K, see Fig. B.5. Each 2-face of K is defined by three vertices. If one considers for each vertex all possible pairs with other vertices, then each 2-face is counted three times. Hence, there are $(d+1)(d-1)d/6$ 2-faces. The dimension of $P_3(K)$ is given by

$$\dim P_3(K) = (d+1) + d(d+1) + \frac{(d-1)d(d+1)}{6} = \frac{(d+1)(d+2)(d+3)}{6}.$$

For the functionals

$$\left\{\begin{aligned} \Phi_i(v) &= v(\boldsymbol{a}_i),\ i=1,\ldots,d+1, && \text{(vertex)},\\ \Phi_{iij}(v) &= v(\boldsymbol{a}_{iij}),\ i,j=1,\ldots,d+1, i\neq j, && \text{(point on edge)},\\ \Phi_{ijk}(v) &= v(\boldsymbol{a}_{ijk}),\ i=1,\ldots,d+1, i<j<k, && \text{(point on 2-face)} \end{aligned}\right\},$$

the local basis is given by

$$\left\{\phi_i(\lambda) = \frac{1}{2}\lambda_i(3\lambda_i - 1)(3\lambda_i - 2), \quad \phi_{iij}(\lambda) = \frac{9}{2}\lambda_i\lambda_j(3\lambda_i - 1), \right.$$

$$\left. \phi_{ijk}(\lambda) = 27\lambda_i\lambda_j\lambda_k \right\}.$$

In two dimensions, the function $\phi_{ijk}(\lambda)$ is called cell bubble function. □

Example B.41 (P_1^{bubble}) The P_1^{bubble} finite element is just the P_1 finite element enriched with mesh cell bubbles. In two dimensions, the functionals are given by the point values of a function $v(\boldsymbol{x})$ in the vertices $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$, and by

$$\frac{3\big[v(\boldsymbol{a}_1) + v(\boldsymbol{a}_2) + v(\boldsymbol{a}_3)\big] + 8\big[v\,(\boldsymbol{a}_{12}) + v\,(\boldsymbol{a}_{13}) + v\,(\boldsymbol{a}_{23})\,\big] + 27v\,(\boldsymbol{a}_{123})}{27},$$

see Figs. B.4 and B.5 for the notations. The corresponding local basis is

$$\{\lambda_1 - 20\lambda_1\lambda_2\lambda_3,\ \lambda_1 - 20\lambda_1\lambda_2\lambda_3,\ \lambda_1 - 20\lambda_1\lambda_2\lambda_3,\ 27\lambda_1\lambda_2\lambda_3\}.$$

□

Example B.42 (P_2^{bubble}) In this space, the P_2 finite element is enriched with bubble functions.

In two dimensions, one can take as nodal functionals the same functionals as for the P_2 element and as seventh functional

$$\frac{3\big[v(\boldsymbol{a}_1) + v(\boldsymbol{a}_2) + v(\boldsymbol{a}_3)\big] + 8\big[v\,(\boldsymbol{a}_{12}) + v\,(\boldsymbol{a}_{13}) + v\,(\boldsymbol{a}_{23})\,\big] + 27v\,(\boldsymbol{a}_{123})}{20},$$

compare Figs. B.4 and B.5 for the notations. Then, the local basis is given by

$$\{4\lambda_1\lambda_2 - 20\lambda_1\lambda_2\lambda_3,\ 4\lambda_1\lambda_3 - 20\lambda_1\lambda_2\lambda_3,\ 4\lambda_2\lambda_3 - 20\lambda_1\lambda_2\lambda_3,$$

$$2\lambda_1(\lambda_1 - 0.5),\ 2\lambda_2(\lambda_2 - 0.5),\ 2\lambda_1(\lambda_2 - 0.5),\ 20\lambda_1\lambda_2\lambda_3\}.$$

In the three-dimensional case, the enrichment is performed with the mesh cell bubble function and with the four bubble functions on the faces. The functionals are the four values in the vertices, the six values on the mid points of the edges, the four values in the barycenters of the faces, and the value in the barycenter of the mesh cell. Altogether, there are 15 functionals. The local basis is given by

$$\{\lambda_1(2\lambda_1 - 1) + 3\lambda_1(\lambda_2\lambda_3 + \lambda_2\lambda_4 + \lambda_3\lambda_4) - 4\lambda_1\lambda_2\lambda_3\lambda_4, \ldots,$$

$$\lambda_1\lambda_2(4 - 12\lambda_4 - 12\lambda_3 + 32\lambda_3\lambda_4), \ldots,$$

$$27\lambda_1\lambda_2\lambda_3(1-4\lambda_4),\ 27\lambda_1\lambda_2(1-4\lambda_3)\lambda_4,\ 27\lambda_1(1-4\lambda_2)\lambda_3\lambda_4,$$
$$27(1-4\lambda_1)\lambda_2\lambda_3\lambda_4,\ 256\lambda_1\lambda_2\lambda_3\lambda_4\big\},$$

where the remaining basis functions are given by appropriate permutations of the indices. □

Example B.43 (P_1^{nc}: Non-conforming Linear Finite Element, Crouzeix–Raviart Finite Element, Crouzeix and Raviart (1973)) This finite element consists of piecewise linear but discontinuous functions. The functionals are given by the values of the functions in the barycenters of the faces such that $\dim P_1^{\mathrm{nc}}(K) = (d+1)$. It follows from the definition of the finite element space, Definition B.25, that the functions from P_1^{nc} are continuous in the barycenter of the faces

$$P_1^{\mathrm{nc}} = \big\{v \in L^2(\Omega) \ : \ v|_K \in P_1(K),\ v(\boldsymbol{x}) \text{ is continuous at the barycenter of all faces}\big\}. \tag{B.19}$$

Equivalently, the functionals can be defined to be the integral mean values on the faces and then the global space is defined to be

$$P_1^{\mathrm{nc}} = \left\{ v \in L^2(\Omega) \ : \ v|_K \in P_1(K), \int_E v|_K \, d\boldsymbol{s} = \int_E v|_{K'} \, d\boldsymbol{s} \ \forall\, E \in \mathcal{E}(K) \cap \mathcal{E}(K') \right\}, \tag{B.20}$$

where $\mathcal{E}(K)$ is the set of all $(d-1)$-dimensional faces of K (Fig. B.6).

For the description of this finite element, one defines the functionals by

$$\Phi_i(v) = v(\boldsymbol{a}_{i-1,i+1}) \text{ for } d=2, \quad \Phi_i(v) = v(\boldsymbol{a}_{i-2,i-1,i+1}) \text{ for } d=3,$$

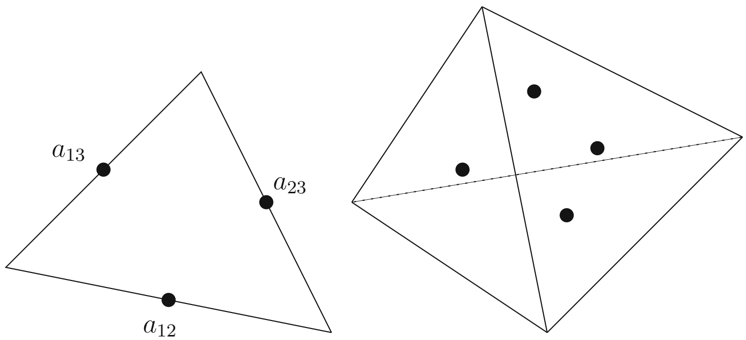

Fig. B.6 The finite element $P_1^{\mathrm{nc}}(K)$

where the points are the barycenters of the faces with the vertices that correspond to the indices. This system is unisolvent with the local basis

$$\phi_i(\lambda) = 1 - d\lambda_i, \quad i = 1, \ldots, d+1.$$

□

Example B.44 (P_1^{disc}) This space consists of piecewise linear but discontinuous functions.

On the reference mesh cell $\hat{K}$ in two dimensions, one can use the functionals applied to a function $v(\hat{\boldsymbol{x}})$ given by

$$\int_{\hat{K}} 2v(\hat{\boldsymbol{x}})\, d\hat{\boldsymbol{x}}, \quad \int_{\hat{K}} (24\hat{x} - 8)\, v(\hat{\boldsymbol{x}})\, d\hat{\boldsymbol{x}}, \quad \int_{\hat{K}} (24\hat{y} - 8)\, v(\hat{\boldsymbol{x}})\, d\hat{\boldsymbol{x}}$$

and the corresponding local basis is

$$\left\{1,\ \hat{\lambda}_2 - \hat{\lambda}_1,\ \hat{\lambda}_3 - \hat{\lambda}_1\right\} = \{1,\ 2\hat{x} + \hat{y} - 1,\ \hat{x} + 2\hat{y} - 1\}.$$

In three dimensions, let $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4$ be the vertices of the tetrahedron and S_K its barycenter. Then, the following functionals can be used

$$\frac{v(\boldsymbol{a}_1) + v(\boldsymbol{a}_2) + v(\boldsymbol{a}_3) + v(\boldsymbol{a}_4) + 16v(S_K)}{120},$$
$$\frac{-v(\boldsymbol{a}_1) + 3v(\boldsymbol{a}_2) - v(\boldsymbol{a}_3) - v(\boldsymbol{a}_4)}{4}, \quad \frac{-v(\boldsymbol{a}_1) - v(\boldsymbol{a}_2) + 3v(\boldsymbol{a}_3) - v(\boldsymbol{a}_4)}{4},$$
$$\frac{-v(\boldsymbol{a}_1) - v(\boldsymbol{a}_2) - v(\boldsymbol{a}_3) + 3v(\boldsymbol{a}_4)}{4}.$$

The corresponding local basis is given by

$$\{6,\ \lambda_2 - \lambda_1,\ \lambda_3 - \lambda_1,\ \lambda_4 - \lambda_1\}.$$

□

Example B.45 (Raviart–Thomas Finite Elements RT_k*)* Raviart–Thomas finite elements are a class of vector-valued finite elements that approximate the space $H(\mathrm{div}, \Omega)$, see (3.37). Details of their definition and important properties can be found, e.g., in Boffi et al. (2013, p. 84).

Consider a simplicial triangulation with mesh cells $\{K\}$ and let $\boldsymbol{P}_k(K) = (P_k(K))^d$, $k \geq 0$. Then, the following local polynomial space is defined directly on K

$$\mathrm{RT}_k(K) = \{\boldsymbol{v} \in (\boldsymbol{P}_k(K) + \boldsymbol{x}P_k(K)) \ :\ \boldsymbol{v} \cdot \boldsymbol{n}_{\partial K} \in R_k(\partial K)\}, \quad k \geq 0,$$

where

$$R_k(\partial K) = \left\{\varphi \in L^2(\partial K) \ : \ \varphi|_E \in P_k(E) \text{ for all faces } E \subset \partial K\right\}.$$

It is noted in Boffi et al. (2013, Remark 2.3.1) that this definition of $\mathrm{RT}_k(K)$ is different than the original definition in Raviart and Thomas (1977).

In particular, the Raviart–Thomas element of lowest order is given by

$$\mathrm{RT}_0(K) = \{\boldsymbol{v} \in (\boldsymbol{P}_0(K) + \boldsymbol{x}P_0(K)) \ : \ \boldsymbol{v} \cdot \boldsymbol{n}_{\partial K} \in R_0(\partial K)\},$$

i.e., it is linear on K. Its local functionals are $\boldsymbol{v} \cdot \boldsymbol{n}_{E_i}$, $i = 1, \ldots, d+1$. A function from $\mathrm{RT}_0(K)$ can be written in the form

$$\boldsymbol{v}(\boldsymbol{x}) = \boldsymbol{a} + b\boldsymbol{x}, \quad \boldsymbol{a} \in \mathbb{R}^d, b \in \mathbb{R}, \boldsymbol{x} \in K. \tag{B.21}$$

A face $E \subset \partial K$ is a hyperplane that can be represented in the form

$$\boldsymbol{x} \cdot \boldsymbol{n}_E = c, \quad c \in \mathbb{R}, \ \forall\, \boldsymbol{x} \in E.$$

Inserting this representation in (B.21) yields

$$\boldsymbol{v}(\boldsymbol{x}) \cdot \boldsymbol{n}_E = \boldsymbol{a} \cdot \boldsymbol{n}_E + b\boldsymbol{x} \cdot \boldsymbol{n}_E = \boldsymbol{a} \cdot \boldsymbol{n}_E + bc = \text{const} \quad \forall\, \boldsymbol{x} \in E.$$

Thus, the normal component of $\boldsymbol{v}$ on each face is a constant.

The global space RT_0 is defined as usual by defining global functionals on the basis of the local functionals and requiring the continuity of the global functionals, see Definition B.25. Consequently, the normal component of functions from RT_0 is continuous across faces of the mesh cells. Since the normal component on each face is a constant, it is sufficient for requiring its continuity to require the continuity in the barycenters $\{\boldsymbol{m}_E\}$ of $\{E\}$. From Lemma 3.66, it follows that $\mathrm{RT}_0 \subset H(\mathrm{div}, \Omega)$. □

B.4 Finite Elements on Parallelepipeds and Quadrilaterals

Remark B.46 (Reference Mesh Cells, Reference Map to Parallelepipeds) One can find in the literature two reference cells: the unit cube $[0,1]^d$ and the large unit cube $[-1,1]^d$. It does not matter which reference cell is chosen. Here, the large unit cube will be used: $\hat{K} = [-1,1]^d$. The class of admissible reference maps $\{F_K\}$ to parallelepipeds consists of bijective affine mappings of the form

$$F_K\hat{\boldsymbol{x}} = B_K\hat{\boldsymbol{x}} + \boldsymbol{b}, \quad B_K \in \mathbb{R}^{d\times d}, \ \boldsymbol{b} \in \mathbb{R}^d.$$

If B_K is a diagonal matrix, then $\hat{K}$ is mapped to d-rectangles.

The class of mesh cells that is obtained in this way is not sufficient to triangulate general domains. If one wants to use more general mesh cells than parallelepipeds, then the class of admissible reference maps has to be enlarged, see Remark B.55. □

Definition B.47 (Polynomial Space Q_k) Let $\boldsymbol{x} = (x_1, \ldots, x_d)^T$ and denote by $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)^T$ a multi-index. Then, the polynomial space Q_k is given by

$$Q_k = \mathrm{span}\left\{ \prod_{i=1}^{d} x_i^{\alpha_i} = \boldsymbol{x}^{\boldsymbol{\alpha}} \; : \; 0 \le \alpha_i \le k \;\text{ for }\; i = 1, \ldots, d \right\} .$$

□

Remark B.48 (Finite Elements on d-Rectangles) For simplicity of presentation, the examples below consider d-rectangles. In this case, the finite elements are just tensor products of one-dimensional finite elements. In particular, the basis functions can be written as products of one-dimensional basis functions. □

Example B.49 (Q_0: Piecewise Constant Finite Element) Similarly to the P_0 space, the space Q_0 consists of piecewise constant, discontinuous functions. The functional is the value of the function in the barycenter of the mesh cell K and it holds $\dim Q_0(K) = 1$. □

Example B.50 (Q_1: Conforming Piecewise d-Linear Finite Element) This finite element space is a subspace of $C(\overline{\Omega})$. The functionals are the values of the function in the vertices of the mesh cell, see Fig. B.7. Hence, it is $\dim Q_1(K) = 2^d$.

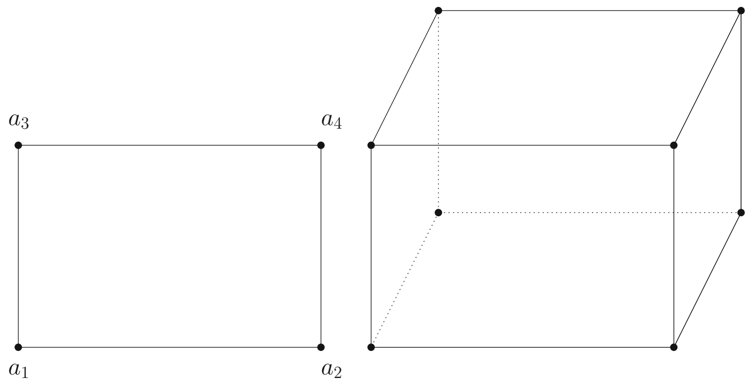

Fig. B.7 The finite element $Q_1(K)$

The one-dimensional local basis functions, which will be used for the tensor product, are given by

$$\hat{\phi}_1(\hat{x}) = \frac{1}{2}(1-\hat{x}), \quad \hat{\phi}_2(\hat{x}) = \frac{1}{2}(1+\hat{x}).$$

With these functions, e.g., the basis functions in two dimensions are computed by

$$\hat{\phi}_1(\hat{x})\hat{\phi}_1(\hat{y}),\ \hat{\phi}_1(\hat{x})\hat{\phi}_2(\hat{y}),\ \hat{\phi}_2(\hat{x})\hat{\phi}_1(\hat{y}),\ \hat{\phi}_2(\hat{x})\hat{\phi}_2(\hat{y}).$$

The continuity of the functions of the finite element space Q_1 is proved in the same way as for simplicial finite elements. It is used that the restriction of a function from $Q_k(K)$ to a face E is a function from the space $Q_k(E)$, $k \geq 1$. □

Example B.51 (Q_2: Conforming Piecewise d-Quadratic Finite Element) It holds that $Q_2 \subset C(\overline{\Omega})$. The functionals in one dimension are the values of the function at both ends of the interval and in the center of the interval, see Fig. B.8. In d dimensions, they are the corresponding values of the tensor product of the intervals. It follows that $\dim Q_2(K) = 3^d$.

The one-dimensional basis function on the reference interval are defined by

$$\hat{\phi}_1(\hat{x}) = -\frac{1}{2}\hat{x}(1-\hat{x}), \quad \hat{\phi}_2(\hat{x}) = (1-\hat{x})(1+\hat{x}), \quad \hat{\phi}_3(\hat{x}) = \frac{1}{2}(1+\hat{x})\hat{x}.$$

The basis function $\prod_{i=1}^{d} \hat{\phi}_2(\hat{x}_i)$ is called cell bubble function. □

Example B.52 (Q_3: Conforming Piecewise d-Cubic Finite Element) This finite element space is a subspace of $C(\overline{\Omega})$. The functionals on the reference interval are given by the values at the end of the interval and the values at the points $\hat{x} = -1/3$, $\hat{x} = 1/3$. In multiple dimensions, it is the corresponding tensor product, see Fig. B.9. The dimension of the local space is $\dim Q_3(K) = 4^d$.

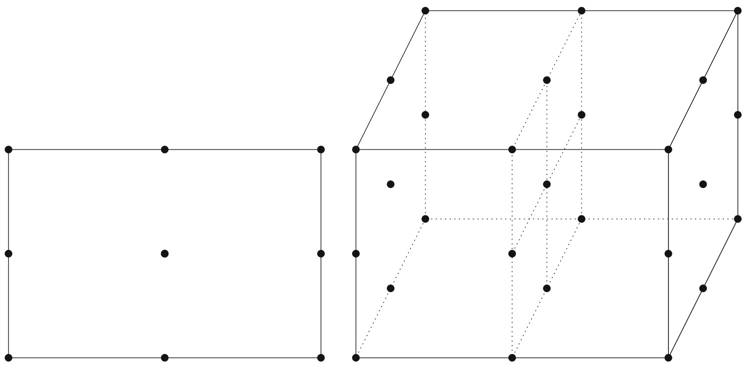

Fig. B.8 The finite element $Q_2(K)$

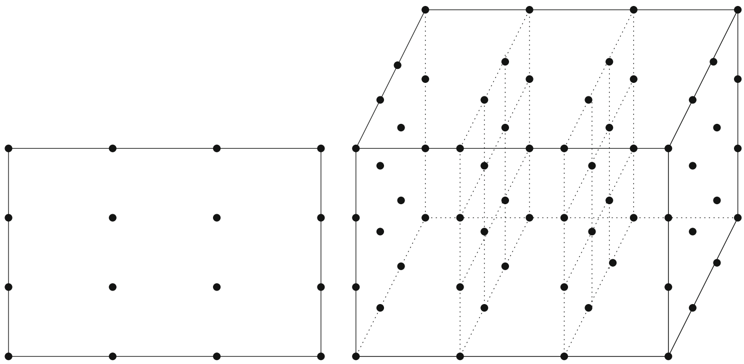

Fig. B.9 The finite element $Q_3(K)$

The one-dimensional basis functions in the reference interval are given by

$$\hat{\phi}_1(\hat{x}) = -\frac{1}{16}(3\hat{x}+1)(3\hat{x}-1)(\hat{x}-1), \quad \hat{\phi}_2(\hat{x}) = \frac{9}{16}(\hat{x}+1)(3\hat{x}-1)(\hat{x}-1),$$
$$\hat{\phi}_3(\hat{x}) = -\frac{9}{16}(\hat{x}+1)(3\hat{x}+1)(\hat{x}-1), \quad \hat{\phi}_4(\hat{x}) = \frac{1}{16}(3\hat{x}+1)(3\hat{x}-1)(\hat{x}+1).$$

□

Example B.53 (Q_1^{rot}: Rotated Non-conforming Element of Lowest Order, Rannacher–Turek Element, Rannacher and Turek (1992)) This finite element space is a generalization of the P_1^{nc} finite element to quadrilateral and hexahedral mesh cells. It consists of discontinuous functions that are continuous at the barycenter of the faces. The dimension of the local finite element space is $\dim Q_1^{\mathrm{rot}}(K) = 2d$. The space on the reference mesh cell is defined by

$$Q_1^{\mathrm{rot}}\left(\hat{K}\right) = \left\{\hat{p} \ : \ \hat{p} \in \mathrm{span}\{1, \hat{x}, \hat{y}, \hat{x}^2 - \hat{y}^2\}\right\} \qquad \text{for } d = 2,$$
$$Q_1^{\mathrm{rot}}\left(\hat{K}\right) = \left\{\hat{p} \ : \ \hat{p} \in \mathrm{span}\{1, \hat{x}, \hat{y}, \hat{z}, \hat{x}^2 - \hat{y}^2, \hat{y}^2 - \hat{z}^2\}\right\} \text{ for } d = 3.$$

Note that the transformed space

$$Q_1^{\mathrm{rot}}(K) = \{p = \hat{p} \circ F_K^{-1}, \hat{p} \in Q_1^{\mathrm{rot}}(\hat{K})\}$$

contains polynomials of the form $ax^2 - by^2$, where a, b depend on F_K.

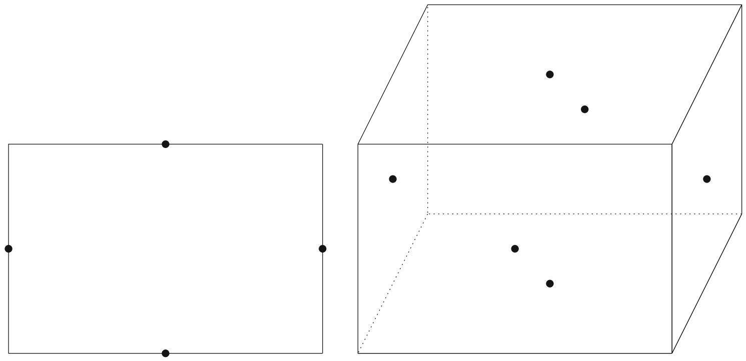

Fig. B.10 The finite element $Q_1^{\text{rot}}(K)$

For $d = 2$, the local basis on the reference cell is given by

$$\hat{\phi}_1(\hat{x},\hat{y}) = -\frac{3}{8}(\hat{x}^2-\hat{y}^2) - \frac{1}{2}\hat{y} + \frac{1}{4}, \quad \hat{\phi}_2(\hat{x},\hat{y}) = \frac{3}{8}(\hat{x}^2-\hat{y}^2) + \frac{1}{2}\hat{x} + \frac{1}{4},$$
$$\hat{\phi}_3(\hat{x},\hat{y}) = -\frac{3}{8}(\hat{x}^2-\hat{y}^2) + \frac{1}{2}\hat{y} + \frac{1}{4}, \quad \hat{\phi}_4(\hat{x},\hat{y}) = \frac{3}{8}(\hat{x}^2-\hat{y}^2) - \frac{1}{2}\hat{x} + \frac{1}{4}. \tag{B.22}$$

Analogously to the Crouzeix–Raviart finite element, the functionals can be defined as point values of the functions in the barycenters of the faces, see Fig. B.10, or as integral mean values of the functions at the faces. Consequently, the finite element spaces are defined in the same way as (B.19) or (B.20), with $P_1^{\text{nc}}(K)$ replaced by $Q_1^{\text{rot}}(K)$.

For a discussion of the practical use of this finite element, it is referred to Remark 3.155. □

Example B.54 (P_k^{disc}, $k \geq 1$) The space P_k^{disc}, $k \geq 1$, is given by

$$P_k^{\text{disc}} = \left\{ v \in L^2(\Omega) \ : \ v|_K \in P_k(K) \right\}.$$

The construction of a basis on the reference mesh cell is based on the Legendre polynomials in $[-1, 1]$, which are given by

$$1, \hat{x}, \frac{1}{2}\left(3\hat{x}^2 - 1\right), \ldots . \tag{B.23}$$

Then, the basis of $P_k^{\text{disc}}(\hat{K})$ in multiple dimensions is defined as a tensor product of polynomials of type (B.23) that gives a polynomial of degree smaller than or equal to k. Concretely, the basis of $P_1^{\mathrm{disc}}(\hat{K})$ is given by

$$\begin{cases} 2d: & \{1, \hat{x}, \hat{y}\}, \\ 3d: & \{1, \hat{x}, \hat{y}, \hat{z}\}, \end{cases} \tag{B.24}$$

and of $P_2^{\mathrm{disc}}(\hat{K})$

$$\begin{cases} 2d: & \left\{1, \hat{x}, \hat{y}, \hat{x}\hat{y}, \frac{1}{2}\left(3\hat{x}^2 - 1\right), \frac{1}{2}\left(3\hat{y}^2 - 1\right)\right\}, \\ 3d: & \left\{1, \hat{x}, \hat{y}, \hat{z}, \hat{x}\hat{y}, \hat{x}\hat{z}, \hat{y}\hat{z}, \frac{1}{2}\left(3\hat{x}^2 - 1\right), \frac{1}{2}\left(3\hat{y}^2 - 1\right), \frac{1}{2}\left(3\hat{z}^2 - 1\right)\right\}. \end{cases} \tag{B.25}$$

For the definition of the local nodal functionals, the $L^2(\hat{K})$ orthogonality of the Legendre polynomials is used. Denoting the basis functions of (B.24) and (B.25) by $\hat{\phi}_j(\hat{\boldsymbol{x}})$, then these functionals are defined by

$$\Phi_{\hat{K},i}\left(\hat{\phi}_j\right) = \left(\int_{\hat{K}} \hat{\phi}_j^2 \, d\hat{\boldsymbol{x}}\right)^{-1} \int_{\hat{K}} \hat{\phi}_i \hat{\phi}_j \, d\hat{\boldsymbol{x}},$$

such that $\Phi_{\hat{K},i}\left(\hat{\phi}_j\right) = \delta_{ij}$. □

Remark B.55 (Parametric Mappings) The image of an affine mapping of the reference mesh cell $\hat{K} = [-1, 1]^d$, $d \in \{2, 3\}$, is a parallelepiped. If one wants to consider finite elements on general d-quadrilaterals, then the class of admissible reference maps has to be enlarged.

The simplest non-affine parametric finite element on quadrilaterals in two dimensions uses bilinear mappings. Let $\hat{K} = [-1, 1]^2$ and let

$$F_K(\hat{\boldsymbol{x}}) = \begin{pmatrix} F_K^1(\hat{\boldsymbol{x}}) \\ F_K^2(\hat{\boldsymbol{x}}) \end{pmatrix} = \begin{pmatrix} a_{11} + a_{12}\hat{x} + a_{13}\hat{y} + a_{14}\hat{x}\hat{y} \\ a_{21} + a_{22}\hat{x} + a_{23}\hat{y} + a_{24}\hat{x}\hat{y} \end{pmatrix}, F_K^i \in Q_1, \ i = 1, 2,$$

be a bilinear mapping from $\hat{K}$ on the class of admissible quadrilaterals. A quadrilateral K is called admissible if

- the length of all edges of K is larger than zero,
- the interior angles of K are smaller than π, i.e., K is convex.

This class contains, e.g., trapezoids and rhombi. □

Remark B.56 (Parametric Finite Element Functions) The functions of the local space $P(K)$ on the mesh cell K are defined by $p = \hat{p} \circ F_K^{-1}$. These functions are in general rational functions. However, using d-linear mappings, then the restriction of F_K on an edge of $\hat{K}$ is an affine map. For instance, in the case of the Q_1 finite element, the functions on K are linear functions on each edge of K. It follows that the functions of the corresponding finite element space are continuous, compare Example B.38. □

B.5 Transform of Integrals

Remark B.57 (Motivation) The transformation of integrals from the reference mesh cell to mesh cells of the grid and vice versa is used as well for the analysis as for the implementation of finite element methods. This section provides an overview of the most important formulae for transformations.

Let $\hat{K} \subset \mathbb{R}^d$ be the reference mesh cell, K be an arbitrary mesh cell, and $F_K : \hat{K} \to K$ with $\boldsymbol{x} = F_K(\hat{\boldsymbol{x}})$ be the reference map. It is assumed that the reference map is a continuous differentiable one-to-one map. The inverse map is denoted by $F_K^{-1} : K \to \hat{K}$. For the integral transforms, the derivatives (Jacobians) of F_K and F_K^{-1} are needed

$$DF_K(\hat{\boldsymbol{x}})_{ij} = \frac{\partial x_i}{\partial \xi_j}, \quad DF_K^{-1}(\boldsymbol{x})_{ij} = \frac{\partial \xi_i}{\partial x_j}, \quad i,j = 1,\ldots,d.$$

□

Remark B.58 (Integral with a Function Without Derivatives) This integral transforms with the standard rule of integral transforms

$$\int_K v(\boldsymbol{x})\, d\boldsymbol{x} = \int_{\hat{K}} \hat{v}(\hat{\boldsymbol{x}})\, |\det DF_K(\hat{\boldsymbol{x}})|\, d\hat{\boldsymbol{x}}, \tag{B.26}$$

where $\hat{v}(\hat{\boldsymbol{x}}) = v(F_K(\hat{\boldsymbol{x}}))$. □

Remark B.59 (Transform of Derivatives) Using the chain rule, one obtains

$$\begin{aligned}\frac{\partial v}{\partial x_i}(\boldsymbol{x}) &= \sum_{j=1}^{d} \frac{\partial \hat{v}}{\partial \xi_j}(\hat{\boldsymbol{x}}) \frac{\partial \xi_j}{\partial x_i} = \nabla_{\hat{\boldsymbol{x}}} \hat{v}(\hat{\boldsymbol{x}}) \cdot \left(\left(DF_K^{-1}(\boldsymbol{x}) \right)^T \right)_i \\ &= \nabla_{\hat{\boldsymbol{x}}} \hat{v}(\hat{\boldsymbol{x}}) \cdot \left(\left(DF_K^{-1}(F_K(\hat{\boldsymbol{x}})) \right)^T \right)_i, \end{aligned} \tag{B.27}$$

$$\begin{aligned}\frac{\partial \hat{v}}{\partial \xi}(\hat{\boldsymbol{x}}) &= \sum_{j=1}^{d} \frac{\partial v}{\partial x_j}(\boldsymbol{x}) \frac{\partial x_j}{\partial \xi_i} = \nabla v(\boldsymbol{x}) \cdot \left((DF_K(\hat{\boldsymbol{x}}))^T \right)_i \\ &= \nabla v(\boldsymbol{x}) \cdot \left(\left(DF_K(F_K^{-1}(\boldsymbol{x})) \right)^T \right)_i. \end{aligned} \tag{B.28}$$

The index i denotes the ith row of a matrix. Derivatives on the reference mesh cell are marked with a symbol on the operator. □

Remark B.60 (Integrals with a Gradients) Using the rule for transforming integrals and (B.27) gives

$$\int_K \boldsymbol{b}(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x})\, d\boldsymbol{x}$$
$$= \int_{\hat{K}} \boldsymbol{b}\left(F_K(\hat{\boldsymbol{x}})\right) \cdot \left[\left(DF_K^{-1}\right)^T \left(F_K(\hat{\boldsymbol{x}})\right)\right] \nabla_{\hat{\boldsymbol{x}}} \hat{v}(\hat{\boldsymbol{x}})\, |\det DF_K(\hat{\boldsymbol{x}})|\; d\hat{\boldsymbol{x}}. \tag{B.29}$$

Similarly, one obtains

$$\int_K \nabla v(\boldsymbol{x}) \cdot \nabla w(\boldsymbol{x})\, d\boldsymbol{x}$$
$$= \int_{\hat{K}} \left[\left(DF_K^{-1}\right)^T \left(F_K(\hat{\boldsymbol{x}})\right)\right] \nabla_{\hat{\boldsymbol{x}}} \hat{v}(\hat{\boldsymbol{x}}) \cdot \left[\left(DF_K^{-1}\right)^T \left(F_K(\hat{\boldsymbol{x}})\right)\right] \nabla_{\hat{\boldsymbol{x}}} \hat{w}(\hat{\boldsymbol{x}})$$
$$\times |\det DF_K(\hat{\boldsymbol{x}})|\; d\hat{\boldsymbol{x}}. \tag{B.30}$$

□

Remark B.61 (Integral with the Divergence) Integrals of the following type are important for the Navier–Stokes equations

$$\int_K \nabla \cdot v(\boldsymbol{x}) q(\boldsymbol{x})\, d\boldsymbol{x} = \int_K \sum_{i=1}^{d} \frac{\partial v_i}{\partial x_i}(\boldsymbol{x}) q(\boldsymbol{x})\, d\boldsymbol{x}$$
$$= \int_{\hat{K}} \sum_{i=1}^{d} \left[\left(\left(DF_K^{-1}(F_K(\hat{\boldsymbol{x}}))\right)^T\right)_i \cdot \nabla_{\hat{\boldsymbol{x}}} \hat{v}_i(\hat{\boldsymbol{x}})\right] \hat{q}(\hat{\boldsymbol{x}})\, |\det DF_K(\hat{\boldsymbol{x}})|\; d\hat{\boldsymbol{x}}$$
$$= \int_{\hat{K}} \left[\left(DF_K^{-1}(F_K(\hat{\boldsymbol{x}}))\right)^T : D_{\hat{\boldsymbol{x}}} \hat{\boldsymbol{v}}(\hat{\boldsymbol{x}})\right] \hat{q}(\hat{\boldsymbol{x}})\, |\det DF_K(\hat{\boldsymbol{x}})|\; d\hat{\boldsymbol{x}}. \tag{B.31}$$

In the derivation, (B.27) was used. □

Example B.62 (Affine Transform) The most important class of reference maps are affine transforms (B.18), where the invertible matrix B_K and the vector $\boldsymbol{b}$ are constants. It follows that

$$\hat{\boldsymbol{x}} = B_K^{-1}\left(\boldsymbol{x} - \boldsymbol{b}\right) = B_K^{-1}\boldsymbol{x} - B_K^{-1}\boldsymbol{b}.$$

In this case, there are

$$DF_K = B_K, \quad DF_K^{-1} = B_K^{-1}, \quad \det DF_K = \det\left(B_K\right).$$

One obtains for the integral transforms from (B.26), (B.29), (B.30), and (B.31)

$$\int_K v(\boldsymbol{x})\, d\boldsymbol{x} = |\det(B_K)| \int_{\hat{K}} \hat{v}(\hat{\boldsymbol{x}})\, d\hat{\boldsymbol{x}}, \tag{B.32}$$

$$\int_K \boldsymbol{b}(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x})\, d\boldsymbol{x} = |\det(B_K)| \int_{\hat{K}} \boldsymbol{b}\left(F_K(\hat{\boldsymbol{x}})\right) \cdot B_K^{-T} \nabla_{\hat{\boldsymbol{x}}} \hat{v}(\hat{\boldsymbol{x}})\, d\hat{\boldsymbol{x}}, \tag{B.33}$$

$$\int_K \nabla v(\boldsymbol{x}) \cdot \nabla w(\boldsymbol{x})\, d\boldsymbol{x} = |\det(B_K)| \int_{\hat{K}} B_K^{-T} \nabla_{\hat{\boldsymbol{x}}} \hat{v}(\hat{\boldsymbol{x}}) \cdot B_K^{-T} \nabla_{\hat{\boldsymbol{x}}} \hat{w}(\hat{\boldsymbol{x}})\, d\hat{\boldsymbol{x}}, \tag{B.34}$$

$$\int_K \nabla \cdot \boldsymbol{v}(\boldsymbol{x}) q(\boldsymbol{x})\, d\boldsymbol{x} = |\det(B_K)| \int_{\hat{K}} \left[B_K^{-T} : D_{\hat{\boldsymbol{x}}} \hat{\boldsymbol{v}}(\hat{\boldsymbol{x}})\right] \hat{q}(\hat{\boldsymbol{x}})\, d\hat{\boldsymbol{x}}. \tag{B.35}$$

Setting $v(\boldsymbol{x}) = 1$ in (B.32) yields

$$|\det(B_K)| = \frac{|K|}{\left|\hat{K}\right|}. \tag{B.36}$$

□

Appendix C
Interpolation

Remark C.1 (Motivation) Variational forms of partial differential equations use functions in Sobolev spaces. The solution of these equations shall be approximated with the Ritz method in finite-dimensional spaces, the finite element spaces. The best possible approximation of an arbitrary function from the Sobolev space by a finite element function is a factor in the upper bound for the finite element error, e.g., see the Lemma of Cea, estimate (B.12).

This section studies the approximation quality of finite element spaces. Estimates are proved for interpolants of functions. Interpolation estimates are of course upper bounds of the best approximation error and they can serve as factors in finite element error estimates. □

C.1 Interpolation in Sobolev Spaces by Polynomials

Lemma C.2 (Unique Determination of a Polynomial with Integral Conditions) *Let Ω be a bounded domain in $\mathbb{R}^d$ with Lipschitz boundary. Let $m \in \mathbb{N} \cup \{0\}$ be given and let for all derivatives with multi-index $\boldsymbol{\alpha}$, $|\boldsymbol{\alpha}| \le m$, a value $a_{\boldsymbol{\alpha}} \in \mathbb{R}$ be prescribed. Then, there is a uniquely determined polynomial $p \in P_m(\Omega)$ such that*

$$\int_\Omega \partial_{\boldsymbol{\alpha}} p(\boldsymbol{x})\, d\boldsymbol{x} = a_{\boldsymbol{\alpha}}, \quad |\boldsymbol{\alpha}| \le m. \tag{C.1}$$

Proof Let $p \in P_m(\Omega)$ be an arbitrary polynomial. It has the form

$$p(\boldsymbol{x}) = \sum_{|\boldsymbol{\beta}| \le m} b_{\boldsymbol{\beta}} \boldsymbol{x}^{\boldsymbol{\beta}}.$$

V. John, *Finite Element Methods for Incompressible Flow Problems*, Springer Series in Computational Mathematics 51, DOI 10.1007/978-3-319-45750-5

Inserting this representation in (C.1) leads to a linear system of equations $M\underline{b} = \underline{a}$ with

$$M = (M_{\boldsymbol{\alpha\beta}}),\ M_{\boldsymbol{\alpha\beta}} = \int_\Omega \partial_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\beta}}\, d\boldsymbol{x},\ \underline{b} = (b_{\boldsymbol{\beta}}),\ \underline{a} = (a_{\boldsymbol{\alpha}}),$$

for $|\boldsymbol{\alpha}|, |\boldsymbol{\beta}| \le m$. Since M is a squared matrix, the linear system of equations possesses a unique solution if and only if M is non-singular.

The proof is performed by contradiction. Assume that M is singular. Then, there exists a non-trivial solution of the homogeneous system. That means, there is a polynomial $q \in P_m(\Omega) \setminus \{0\}$ with

$$\int_\Omega \partial_{\boldsymbol{\alpha}} q(\boldsymbol{x})\, d\boldsymbol{x} = 0 \text{ for all } |\boldsymbol{\alpha}| \le m.$$

The polynomial $q(\boldsymbol{x})$ has the representation $q(\boldsymbol{x}) = \sum_{|\boldsymbol{\beta}|\le m} c_{\boldsymbol{\beta}} \boldsymbol{x}^{\boldsymbol{\beta}}$. Now, one can choose a $c_{\boldsymbol{\beta}} \ne 0$ with maximal value $|\boldsymbol{\beta}|$. Then, it is $\partial_{\boldsymbol{\beta}} q(\boldsymbol{x}) = Cc_{\boldsymbol{\beta}} = const \ne 0$, where $C > 0$ comes from the differentiation rule for polynomials, which is a contradiction to the vanishing of the integral for $\partial_{\boldsymbol{\beta}} q(\boldsymbol{x})$. ■

Lemma C.3 (Poincaré-Type Inequality) *Denote by $D^k v(\boldsymbol{x})$, $k \in \mathbb{N} \cup \{0\}$, the total derivative of order k of a function $v(\boldsymbol{x})$, e.g., for $k = 1$ the gradient of $v(\boldsymbol{x})$. Let Ω be convex and be included into a ball of radius R. Let $l \in \mathbb{N} \cup \{0\}$ with $k \le l$ and let $p \in \mathbb{R}$ with $p \in [1, \infty)$. Assume that $v \in W^{l,p}(\Omega)$ satisfies*

$$\int_\Omega \partial_{\boldsymbol{\alpha}} v(\boldsymbol{x})\, d\boldsymbol{x} = 0 \textit{ for all } |\boldsymbol{\alpha}| \le l - 1,$$

then it holds the estimate

$$\left\| D^k v \right\|_{L^p(\Omega)} \le C R^{l-k} \left\| D^l v \right\|_{L^p(\Omega)},$$

where the constant C does not depend on Ω and on $v(\boldsymbol{x})$.

Proof There is nothing to prove if $k = l$. In addition, it suffices to prove the lemma for $k = 0$ and $l = 1$, since the general case follows by applying the result to $\partial_{\boldsymbol{\alpha}} v(\boldsymbol{x})$.

Since Ω is assumed to be convex, the integral mean value theorem can be written in the form

$$v(\boldsymbol{x}) - v(\boldsymbol{y}) = \int_0^1 \nabla v(t\boldsymbol{x} + (1-t)\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y})\, dt, \quad \boldsymbol{x}, \boldsymbol{y} \in \Omega.$$

Integration with respect to $\boldsymbol{y}$ yields

$$v(\boldsymbol{x}) \int_\Omega d\boldsymbol{y} - \int_\Omega v(\boldsymbol{y})\, d\boldsymbol{y} = \int_\Omega \int_0^1 \nabla v(t\boldsymbol{x} + (1-t)\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y})\, dt\, d\boldsymbol{y}.$$

It follows from the assumption that the second integral on the left-hand side vanishes that

$$v(\boldsymbol{x}) = \frac{1}{|\Omega|}\int_\Omega \int_0^1 \nabla v(t\boldsymbol{x} + (1-t)\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y})\, dt\, d\boldsymbol{y}.$$

Now, taking the absolute value on both sides, using that the absolute value of an integral is estimated from above by the integral of the absolute value, applying the Cauchy–Schwarz inequality for vectors (A.2), and the estimate $\|\boldsymbol{x}-\boldsymbol{y}\|_2 \le 2R$ yields

$$\begin{aligned} |v(\boldsymbol{x})| &= \frac{1}{|\Omega|}\left|\int_\Omega \int_0^1 \nabla v(t\boldsymbol{x} + (1-t)\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y})\, dt\, d\boldsymbol{y}\right| \\ &\le \frac{1}{|\Omega|}\int_\Omega \int_0^1 |\nabla v(t\boldsymbol{x} + (1-t)\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y})|\, dt\, d\boldsymbol{y} \\ &\le \frac{2R}{|\Omega|}\int_\Omega \int_0^1 \|\nabla v(t\boldsymbol{x} + (1-t)\boldsymbol{y})\|_2\, dt\, d\boldsymbol{y}. \end{aligned} \tag{C.2}$$

Then, (C.2) is raised to the power p and integrated with respect to $\boldsymbol{x}$. One obtains with Hölder's inequality (A.9), with $p^{-1} + q^{-1} = 1 \implies p/q - p = p(1/q - 1) = -1$, that

$$\begin{aligned} \int_\Omega |v(\boldsymbol{x})|^p\, d\boldsymbol{x} &\le \frac{CR^p}{|\Omega|^p}\int_\Omega \left(\int_\Omega \int_0^1 \|\nabla v(t\boldsymbol{x} + (1-t)\boldsymbol{y})\|_2\, dt\, d\boldsymbol{y}\right)^p d\boldsymbol{x} \\ &\le \frac{CR^p}{|\Omega|^p}\int_\Omega \Bigg[\underbrace{\left(\int_\Omega \int_0^1 1^q\, dt\, d\boldsymbol{y}\right)^{p/q}}_{|\Omega|^{p/q}} \\ &\qquad \times \left(\int_\Omega \int_0^1 \|\nabla v(t\boldsymbol{x} + (1-t)\boldsymbol{y})\|_2^p\, dt\, d\boldsymbol{y}\right)\Bigg] d\boldsymbol{x} \\ &= \frac{CR^p}{|\Omega|}\int_\Omega \left(\int_\Omega \int_0^1 \|\nabla v(t\boldsymbol{x} + (1-t)\boldsymbol{y})\|_2^p\, dt\, d\boldsymbol{y}\right) d\boldsymbol{x}. \end{aligned}$$

Applying the theorem of Fubini allows the commutation of the integration

$$\int_\Omega |v(\boldsymbol{x})|^p\, d\boldsymbol{x} \le \frac{CR^p}{|\Omega|}\int_0^1 \int_\Omega \left(\int_\Omega \|\nabla v(t\boldsymbol{x} + (1-t)\boldsymbol{y})\|_2^p\, d\boldsymbol{y}\right) d\boldsymbol{x}\, dt.$$

Using the integral mean value theorem in one dimension gives that there is a $t_0 \in [0, 1]$ such that

$$\int_\Omega |v(\boldsymbol{x})|^p \, d\boldsymbol{x} \le \frac{CR^p}{|\Omega|} \int_\Omega \left(\int_\Omega \|\nabla v(t_0\boldsymbol{x} + (1-t_0)\boldsymbol{y})\|_2^p \, d\boldsymbol{y} \right) d\boldsymbol{x}.$$

The function $\|\nabla v(\boldsymbol{x})\|_2^p$ will be extended to $\mathbb{R}^d$ by zero and the extension will be also denoted by $\|\nabla v(\boldsymbol{x})\|_2^p$. Then, it is

$$\int_\Omega |v(\boldsymbol{x})|^p \, d\boldsymbol{x} \le \frac{CR^p}{|\Omega|} \int_\Omega \left(\int_{\mathbb{R}^d} \|\nabla v(t_0\boldsymbol{x} + (1-t_0)\boldsymbol{y})\|_2^p \, d\boldsymbol{y} \right) d\boldsymbol{x}. \tag{C.3}$$

Let $t_0 \in [0, 1/2]$. Since the domain of integration is $\mathbb{R}^d$, a substitution of variables $t_0\boldsymbol{x} + (1-t_0)\boldsymbol{y} = \boldsymbol{z}$ can be applied and leads to

$$\int_{\mathbb{R}^d} \|\nabla v(t_0\boldsymbol{x} + (1-t_0)\boldsymbol{y})\|_2^p \, d\boldsymbol{y} = \frac{1}{1-t_0} \int_{\mathbb{R}^d} \|\nabla v(\boldsymbol{z})\|_2^p \, d\boldsymbol{z} \le 2 \|\nabla v\|_{L^p(\Omega)}^p ,$$

since $1/(1-t_0) \le 2$. Inserting this expression in (C.3) gives

$$\int_\Omega |v(\boldsymbol{x})|^p \, d\boldsymbol{x} \le 2CR^p \|\nabla v\|_{L^p(\Omega)}^p .$$

If $t_0 > 1/2$ then one changes the roles of $\boldsymbol{x}$ and $\boldsymbol{y}$, applies the theorem of Fubini to change the sequence of integration, and uses the same arguments. ■

Remark C.4 (On Lemma C.3) Lemma C.3 proves an inequality of Poincaré-type. It says that it is possible to estimate the $L^p(\Omega)$ norm of a lower derivative of a function $v(\boldsymbol{x})$ by the same norm of a higher derivative if the integral mean values of some lower derivatives vanish.

An important application of Lemma C.3 is in the proof of the Bramble–Hilbert lemma. The Bramble–Hilbert lemma considers a continuous linear functional that is defined on a Sobolev space and that vanishes for all polynomials of degree less than or equal to m. It states that the value of the functional can be estimated by the Lebesgue norm of the $(m+1)$th total derivative of the functions from this Sobolev space. □

Theorem C.5 (Bramble–Hilbert Lemma) *Let $m \in \mathbb{N} \cup \{0\}$, $m \ge 0$, $p \in [1, \infty]$, and $F : W^{m+1,p}(\Omega) \to \mathbb{R}$ be a continuous linear functional, and let the conditions of Lemmas C.2 and C.3 be satisfied. Let*

$$F(p) = 0 \quad \forall \ p \in P_m(\Omega),$$

then there is a constant $C(\Omega)$, which is independent of v and F, such that

$$|F(v)| \le C(\Omega) \left\| D^{m+1} v \right\|_{L^p(\Omega)} \quad \forall \ v \in W^{m+1,p}(\Omega).$$

Proof Let $v \in W^{m+1,p}(\Omega)$. It follows from Lemma C.2 that there is a polynomial from $P_m(\Omega)$ with

$$\int_\Omega \partial_{\boldsymbol{\alpha}}(v+p)(\boldsymbol{x})\,d\boldsymbol{x} = 0 \text{ for } |\boldsymbol{\alpha}| \le m.$$

Lemma C.3 gives, with $l = m+1$ and considering each term in $\|\cdot\|_{W^{m+1,p}(\Omega)}$ individually, the estimate

$$\|v+p\|_{W^{m+1,p}(\Omega)} \le C(\Omega)\left\|D^{m+1}(v+p)\right\|_{L^p(\Omega)} = C(\Omega)\left\|D^{m+1}v\right\|_{L^p(\Omega)}.$$

From the vanishing of F for $p \in P_m(\Omega)$ and the continuity of F, it follows that

$$|F(v)| = |F(v+p)| \le c\,\|v+p\|_{W^{m+1,p}(\Omega)} \le C(\Omega)\left\|D^{m+1}v\right\|_{L^p(\Omega)}.$$

■

Remark C.6 (Strategy for Estimating the Interpolation Error) The Bramble–Hilbert lemma will be used for estimating the interpolation error for finite elements. The strategy is as follows:

- Show first the estimate on the reference mesh cell $\hat{K}$.
- Transform the estimate on an arbitrary mesh cell K to the reference mesh cell $\hat{K}$.
- Apply the estimate on $\hat{K}$.
- Transform back to K.

One has to study what happens if the transforms are applied to the estimate. □

Remark C.7 (Assumptions, Definition of the Interpolant) Let $\hat{K} \subset \mathbb{R}^d, d \in \{2,3\}$, be a reference mesh cell (compact polyhedron), $\hat{P}(\hat{K})$ a polynomial space of dimension N, and $\hat{\Phi}_1,\ldots,\hat{\Phi}_N : C^s(\hat{K}) \to \mathbb{R}$ continuous linear functionals. It will be assumed that the space $\hat{P}(\hat{K})$ is unisolvent with respect to these functionals. Then, there is a local basis $\hat{\phi}_1,\ldots,\hat{\phi}_N \in \hat{P}(\hat{K})$.

Consider $\hat{v} \in C^s(\hat{K})$, then the interpolant $I_{\hat{K}}\hat{v} \in \hat{P}(\hat{K})$ is defined by

$$I_{\hat{K}}\hat{v}(\hat{\boldsymbol{x}}) = \sum_{i=1}^{N} \hat{\Phi}_i(\hat{v})\hat{\phi}_i(\hat{\boldsymbol{x}}).$$

The operator $I_{\hat{K}}$ is a continuous and linear operator from $C^s(\hat{K})$ to $\hat{P}(\hat{K})$. From the linearity, it follows that $I_{\hat{K}}$ is the identity on $\hat{P}(\hat{K})$

$$I_{\hat{K}}\hat{p} = \hat{p} \quad \forall\, \hat{p} \in \hat{P}(\hat{K}).$$

□

Theorem C.8 (Interpolation Error Estimate on a Reference Mesh Cell) *Let $P_m(\hat{K}) \subset \hat{P}(\hat{K})$, let $p \in [1, \infty)$, and let $\hat{s} \in \mathbb{N} \cup \{0\}$ such that $(m+1-\hat{s})p > d \geq (m-\hat{s})p$ and $\hat{s} \geq s$, where s appears in the definition of the interpolation operator. Then there is a constant C that is independent of $\hat{v}(\hat{\boldsymbol{x}})$ such that*

$$\left\| \hat{v} - I_{\hat{K}}\hat{v} \right\|_{W^{m+1,p}(\hat{K})} \leq C \left\| D^{m+1}\hat{v} \right\|_{L^p(\hat{K})} \quad \forall\ \hat{v} \in W^{m+1,p}(\hat{K}). \tag{C.4}$$

Proof Since $\hat{K}$ is a bounded, one has the Sobolev imbedding (A.20)

$$W^{m+1,p}(\hat{K}) = W^{(m+1-\hat{s})+\hat{s},p}(\hat{K}) \to C^{\hat{s}}(\hat{K}).$$

Because $\hat{K}$ is convex, the imbedding $C^{\hat{s}}(\hat{K}) \to C^{s}(\hat{K})$ is compact, see Adams (1975, Theorem 1.31), such that the interpolation operator is well defined in $W^{m+1,p}(\hat{K})$. From the identity of the interpolation operator in $P_m(\hat{K})$, the triangle inequality, the boundedness of the interpolation operator (it is a linear and continuous operator mapping $C^s(\hat{K}) \to \hat{P}(\hat{K}) \subset W^{m+1,p}(\hat{K})$), and the Sobolev imbedding, one obtains for $\hat{q} \in P_m(\hat{K})$

$$\begin{aligned}
\left\| \hat{v} - I_{\hat{K}}\hat{v} \right\|_{W^{m+1,p}(\hat{K})} &= \left\| \hat{v} + \hat{q} - I_{\hat{K}}(\hat{v} + \hat{q}) \right\|_{W^{m+1,p}(\hat{K})} \\
&\leq \|\hat{v} + \hat{q}\|_{W^{m+1,p}(\hat{K})} + \left\| I_{\hat{K}}(\hat{v} + \hat{q}) \right\|_{W^{m+1,p}(\hat{K})} \\
&\leq \|\hat{v} + \hat{q}\|_{W^{m+1,p}(\hat{K})} + c \, \|\hat{v} + \hat{q}\|_{C^s(\hat{K})} \\
&\leq c \, \|\hat{v} + \hat{q}\|_{W^{m+1,p}(\hat{K})} \, .
\end{aligned}$$

Now, $\hat{q}(\hat{\boldsymbol{x}})$ is chosen such that

$$\int_{\hat{K}} \partial_{\boldsymbol{\alpha}}(\hat{v} + \hat{q}) \, d\hat{\boldsymbol{x}} = 0 \quad \forall \ |\boldsymbol{\alpha}| \leq m$$

holds. Hence, the assumptions of Lemma C.3 are satisfied. It follows that

$$\|\hat{v} + \hat{q}\|_{W^{m+1,p}(\hat{K})} \leq c \left\| D^{m+1}(\hat{v} + \hat{q}) \right\|_{L^p(\hat{K})} = c \left\| D^{m+1}\hat{v} \right\|_{L^p(\hat{K})} \, .$$

■

Definition C.9 (Quasi-Uniform and Regular Family of Triangulations, Brenner and Scott (2008, Definition 4.4.13)) Let $\{\mathcal{T}^h\}$ with $0 < h \leq 1$, be a family of triangulations such that

$$\max_{K \in \mathcal{T}^h} h_K \leq h \, \mathrm{diam}(\Omega),$$

where h_K is the diameter of $K = F_K(\hat{K})$, i.e., the largest distance of two points that are contained in K. The family is called to be quasi-uniform, if there exists a $C > 0$

such that

$$\min_{K\in\mathcal{T}^h} \rho_K \geq Ch \operatorname{diam}(\Omega) \tag{C.5}$$

for all $h \in (0, 1]$, where ρ_K is the diameter of the largest ball contained in K.

The family is called to be regular, if there is exists a $C > 0$ such that for all $K \in \mathcal{T}^h$ and for all $h \in (0, 1]$

$$\rho_K \geq Ch_K.$$

□

Remark C.10 (Euler's Formula for a Simple Closed Polygon) Consider an admissible triangulation of a simply connected domain with polygonal boundary with triangles. Then, there is a relation of the numbers of mesh cells $\{K\}$, vertices $\{V\}$, and edges $\{E\}$, which is called Euler's formula

$$\#K \; - \#E \; + \#V \; = 1,$$

where # denotes the cardinality of the set. □

Remark C.11 (Assumptions on the Reference Mapping and the Triangulation) For deriving the interpolation error estimate for arbitrary mesh cells K, and finally for the finite element space, one has to study the properties of the affine mapping from K to $\hat{K}$ and of the inverse mapping. Here, only the case of an affine family of finite elements whose mesh cells are generated by affine mappings

$$F_K\hat{\boldsymbol{x}} = B_K\hat{\boldsymbol{x}} + \boldsymbol{b},$$

will be considered, see (B.18), where B_K is a non-singular $d \times d$ matrix and $\boldsymbol{b}$ is a d vector.

For the global estimate, a quasi-uniform family of triangulations will be considered. □

Lemma C.12 (Estimates of Matrix Norms) *For each matrix norm $\|\cdot\|$, one has the estimates*

$$\|B_K\| \leq ch_K, \quad \left\|B_K^{-1}\right\| \leq ch_K^{-1}, \tag{C.6}$$

where the constants depend on the matrix norm and on C_R.

Proof Since $\hat{K}$ is a Lipschitz domain with polyhedral boundary, it contains a ball $B(\hat{\boldsymbol{x}}_0, r)$ with $\hat{\boldsymbol{x}}_0 \in \hat{K}$ and some $r > 0$. Hence, $\hat{\boldsymbol{x}}_0 + \hat{\boldsymbol{y}} \in \hat{K}$ for all $\|\hat{\boldsymbol{y}}\|_2 = r$. It follows that the images

$$\boldsymbol{x}_0 = B_K\hat{\boldsymbol{x}}_0 + \boldsymbol{b}, \quad \boldsymbol{x} = B_K(\hat{\boldsymbol{x}}_0 + \hat{\boldsymbol{y}}) + \boldsymbol{b} = \boldsymbol{x}_0 + B_K\hat{\boldsymbol{y}}$$

are contained in K. Since the triangulation is assumed to be quasi-uniform, one obtains for all $\hat{\boldsymbol{y}}$

$$\|B_K\hat{\boldsymbol{y}}\|_2 = \|\boldsymbol{x} - \boldsymbol{x}_0\|_2 \le C_R h_K.$$

Now, it holds for the spectral norm that

$$\|B_K\|_2 = \sup_{\hat{z}\neq \boldsymbol{0}} \frac{\|B_K\hat{z}\|_2}{\|\hat{z}\|_2} = \frac{1}{r} \sup_{\|\hat{z}\|_2=r} \|B_K\hat{z}\|_2 \le \frac{C_R}{r} h_K.$$

A bound of this form, with a possible different constant, holds also for all other matrix norms since all matrix norms are equivalent, see Remark A.8.

The estimate for $\left\|B_K^{-1}\right\|$ proceeds in the same way with interchanging the roles of K and $\hat{K}$. ■

Theorem C.13 (Local Interpolation Estimate) *Let an affine family of finite elements be given by its reference cell* $\hat{K}$*, the functionals* $\{\hat{\Phi}_i\}$*, and a space of polynomials* $\hat{P}(\hat{K})$*. Let all assumptions of Theorem C.8 be satisfied. Then, for all* $v \in W^{m+1,p}(K)$, $p \in [1, \infty)$*, there is a constant* C*, which is independent of* v*, such that*

$$\left\|D^k(v - I_K v)\right\|_{L^p(K)} \le C h_K^{m+1-k} \left\|D^{m+1} v\right\|_{L^p(K)}, \quad 0 \le k \le m+1. \tag{C.7}$$

Proof The idea of the proof consists in transforming the left-hand side of (C.7) to the reference cell, using the interpolation estimate on the reference cell, and transforming back.

(i) Denote the elements of the matrices B_K and B_K^{-1} by b_{ij} and $b_{ij}^{(-1)}$, respectively. Since $\|B_K\|_\infty = \max_{i,j}\left|b_{ij}\right|$ is also a matrix norm, it holds that

$$\left|b_{ij}\right| \le C h_K, \quad \left|b_{ij}^{(-1)}\right| \le C h_K^{-1}. \tag{C.8}$$

Using element-wise estimates for the matrix B_K (Leibniz formula for determinants), one obtains

$$\left|\det B_K\right| \le C h_K^d, \quad \left|\det B_K^{-1}\right| \le C h_K^{-d}. \tag{C.9}$$

(ii) The next step consists in proving that the transformed interpolation operator is equal to the natural interpolation operator on K. The latter one is given by

$$I_K v = \sum_{i=1}^{N} \Phi_{K,i}(v)\phi_{K,i}, \tag{C.10}$$

where $\{\phi_{K,i}\}$ is the basis of the space

$$P(K) = \left\{p \ : \ K \to \mathbb{R} \ : \ p = \hat{p} \circ F_K^{-1}, \hat{p} \in \hat{P}(\hat{K})\right\},$$

which satisfies $\Phi_{K,i}(\phi_{K,j}) = \delta_{ij}$. The functionals are defined by

$$\Phi_{K,i}(v) = \hat{\Phi}_i(v \circ F_K).$$

Hence, it follows with $v = \hat{\phi}_j \circ F_K^{-1}$ from the condition on the local basis on $\hat{K}$ that

$$\Phi_{K,i}(\hat{\phi}_j \circ F_K^{-1}) = \hat{\Phi}_i(\hat{\phi}_j) = \delta_{ij},$$

i.e., the local basis on K is given by $\phi_{K,j} = \hat{\phi}_j \circ F_K^{-1}$. Using (C.10), one gets

$$\begin{aligned} I_{\hat{K}}\hat{v} &= \sum_{i=1}^N \hat{\Phi}_i(\hat{v})\hat{\phi}_i = \sum_{i=1}^N \Phi_{K,i}(\underbrace{\hat{v} \circ F_K^{-1}}_{=v})\, \phi_{K,i} \circ F_K = \left(\sum_{i=1}^N \Phi_{K,i}(v)\phi_{K,i}\right) \circ F_K \\ &= I_K v \circ F_K. \end{aligned}$$

Consequently, $I_{\hat{K}}\hat{v}$ is transformed correctly.

(iii) One obtains with the chain rule

$$\frac{\partial v(\boldsymbol{x})}{\partial \boldsymbol{x}_i} = \sum_{j=1}^d \frac{\partial \hat{v}(\hat{\boldsymbol{x}})}{\partial \hat{\boldsymbol{x}}_j} b_{ji}^{(-1)}, \quad \frac{\partial \hat{v}(\hat{\boldsymbol{x}})}{\partial \hat{\boldsymbol{x}}_i} = \sum_{j=1}^d \frac{\partial v(\boldsymbol{x})}{\partial \boldsymbol{x}_j} b_{ji}.$$

It follows with (C.8) that (with each derivative one obtains an additional factor of B_K or B_K^{-1}, respectively)

$$\left\|D_{\boldsymbol{x}}^k v(\boldsymbol{x})\right\|_2 \le Ch_K^{-k} \left\|D_{\hat{\boldsymbol{x}}}^k \hat{v}(\hat{\boldsymbol{x}})\right\|_2, \quad \left\|D_{\hat{\boldsymbol{x}}}^k \hat{v}(\hat{\boldsymbol{x}})\right\|_2 \le Ch_K^k \left\|D_{\boldsymbol{x}}^k v(\boldsymbol{x})\right\|_2.$$

One gets with (C.9)

$$\int_K \left\|D_{\boldsymbol{x}}^k v(\boldsymbol{x})\right\|_2^p d\boldsymbol{x} \le Ch_K^{-kp} \left|\det B_K\right| \int_{\hat{K}} \left\|D_{\hat{\boldsymbol{x}}}^k \hat{v}(\hat{\boldsymbol{x}})\right\|_2^p d\hat{\boldsymbol{x}} \le Ch_K^{-kp+d} \int_{\hat{K}} \left\|D_{\hat{\boldsymbol{x}}}^k \hat{v}(\hat{\boldsymbol{x}})\right\|_2^p d\hat{\boldsymbol{x}} \tag{C.11}$$

and

$$\int_{\hat{K}} \left\|D_{\hat{\boldsymbol{x}}}^k \hat{v}(\hat{\boldsymbol{x}})\right\|_2^p d\hat{\boldsymbol{x}} \le Ch_K^{kp} \left|\det B_K^{-1}\right| \int_K \left\|D_{\boldsymbol{x}}^k v(\boldsymbol{x})\right\|_2^p d\boldsymbol{x} \le Ch_K^{kp-d} \int_K \left\|D_{\boldsymbol{x}}^k v(\boldsymbol{x})\right\|_2^p d\boldsymbol{x}. \tag{C.12}$$

Using now the interpolation estimate on the reference cell (C.4) yields

$$\left\|D_{\hat{\boldsymbol{x}}}^{k}(\hat{v}-I_{\hat{K}}\hat{v})\right\|_{L^p(\hat{K})}^{p} \le C\left\|D_{\hat{\boldsymbol{x}}}^{m+1}\hat{v}\right\|_{L^p(\hat{K})}^{p}, \quad 0\le k\le m+1. \tag{C.13}$$

It follows that

$$\begin{aligned}\left\|D_{\boldsymbol{x}}^{k}(v-I_{K}v)\right\|_{L^p(K)}^{p} &\le Ch_K^{-kp+d}\left\|D_{\hat{\boldsymbol{x}}}^{k}(\hat{v}-I_{\hat{K}}\hat{v})\right\|_{L^p(\hat{K})}^{p}\\ &\le Ch_K^{-kp+d}\left\|D_{\hat{\boldsymbol{x}}}^{m+1}\hat{v}\right\|_{L^p(\hat{K})}^{p}\\ &\le Ch_K^{(m+1-k)p}\left\|D_{\boldsymbol{x}}^{m+1}v\right\|_{L^p(K)}^{p}.\end{aligned}$$

Taking the pth root proves the statement of the theorem. ■

Remark C.14 (On Estimate (C.7))

- Note that the power of h_K does not depend on p and d.
- Consider a quasi-uniform triangulation and define

$$h = \max_{K\in\mathcal{T}^h}\{h_K\}.$$

  Then, one obtains by summing over all mesh cells an interpolation estimate for the global finite element space

$$\begin{aligned}\left\|D^{k}(v-I^{h}v)\right\|_{L^p(\Omega)} &= \left(\sum_{K\in\mathcal{T}^h}\left\|D^{k}(v-I_{K}v)\right\|_{L^p(K)}^{p}\right)^{1/p}\\ &\le \left(\sum_{K\in\mathcal{T}^h} ch_K^{(m+1-k)p}\left\|D^{m+1}v\right\|_{L^p(K)}^{p}\right)^{1/p}\\ &\le ch^{(m+1-k)}\left\|D^{m+1}v\right\|_{L^p(\Omega)}.\end{aligned} \tag{C.14}$$

□

Corollary C.15 (Interpolation Estimate for Faces) *Let the assumptions of Theorem C.13 be satisfied. Let $K\in\mathcal{T}^h$ be a mesh cell with the face $E\subset\partial K$, then it holds*

$$\left\|v-I_{K}v\right\|_{L^p(E)} \le Ch_K^{(m+1)-1/p}\left\|D^{m+1}v\right\|_{L^p(K)}. \tag{C.15}$$

Proof The idea of the proof consists in transforming the norm on the face to the norm on a face of the reference cell, applying the trace theorem, and applying the back transform for the reference mesh cell.

Performing the transform of the integral on E to a face $\hat{E}\subset\partial\hat{K}$ gives as factor the determinant of the transform. As the determinant of the transform of an integral from K to $\hat{K}$ is proportional to the volume of K, see (C.9), the determinant of the transform of an integral from E to $\hat{E}$ is proportional to the area of E, which is $|E| = Ch_E^{d-1}$. Thus, one gets

$$\|v - I_K v\|_{L^p(E)}^p \le Ch_E^{d-1} \left\|\hat{v} - I_{\hat{K}}\hat{v}\right\|_{L^p(\hat{E})}^p \le Ch_E^{d-1} \left\|\hat{v} - I_{\hat{K}}\hat{v}\right\|_{L^p(\partial\hat{K})}^p .$$

With the trace theorem, Theorem A.34, one obtains

$$\|v - I_K v\|_{L^p(\partial\hat{K})}^p \le Ch_E^{d-1} \left(\left\|\hat{v} - I_{\hat{K}}\hat{v}\right\|_{L^p(\hat{K})}^p + \left\|\nabla\left(\hat{v} - I_{\hat{K}}\hat{v}\right)\right\|_{L^p(\hat{K})}^p \right).$$

Applying the interpolation estimate on the reference cell (C.4), compare (C.13), the transform from $\hat{K}$ to K, see (C.12), and $h_E \le h_K$ yields for $m \ge 0$

$$\begin{aligned} \|v - I_K v\|_{L^p(E)}^p &\le Ch_E^{d-1} \left\|D_{\hat{\boldsymbol{x}}}^{m+1}\hat{v}\right\|_{L^p(\hat{K})}^p \\ &\le Ch_E^{d-1} h_K^{(m+1)p-d} \left\|D_{\boldsymbol{x}}^{m+1} v\right\|_{L^p(K)}^p \le Ch_K^{(m+1)p-1} \left\|D_{\boldsymbol{x}}^{m+1} v\right\|_{L^p(K)}^p . \end{aligned}$$

Taking the pth root proves (C.15). ■

C.2 Interpolation of Non-smooth Functions

Remark C.16 (Motivation) The interpolation theory of Sect. C.1 requires that the interpolation operator is continuous on the Sobolev space to which the function belongs that should be interpolated. But if, e.g., discontinuous functions should be interpolated with continuous, piecewise linear functions, then Sect. C.1 does not provide estimates.

There are two often used interpolation operators for non-smooth functions. The interpolation operator of Clément (1975) is defined for functions from $L^1(\Omega)$ and it can be generalized to more or less all finite elements. The interpolation operator of Scott and Zhang (1990) is more special. It has the advantage that it preserves homogeneous Dirichlet boundary conditions in a natural way. For the Clément interpolation operator, one needs a modification for the preservation of homogeneous Dirichlet boundary conditions, which cannot be generalized easily to the non-homogeneous case. Here, only the interpolation operator of Clément, for linear finite elements, will be considered.

Let $\mathcal{T}^h$ be a regular triangulation of the polyhedral domain $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, with simplices K. Denote by P_1 the space of continuous, piecewise linear finite elements on $\mathcal{T}^h$. □

Remark C.17 (Construction of the Interpolation Operator of Clément) For each vertex V_i of the triangulation, the union of all grid cells that possess V_i as vertex will be denoted by ω_i, see Fig. C.1.

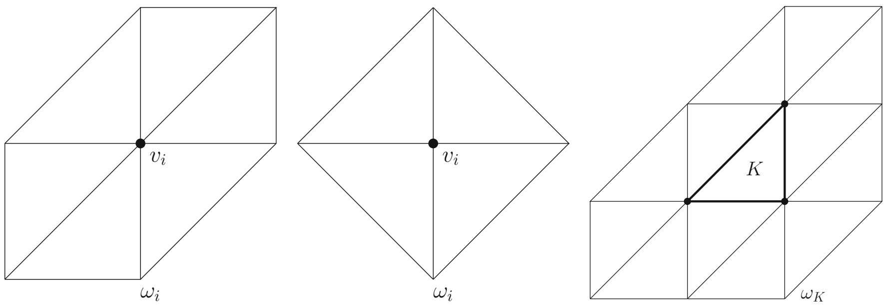

Fig. C.1 Subdomains ω_i (*left and center*) and a subdomain ω_K (*right*)

Let $v \in L^1(\Omega)$ and let $P_1(\omega_i)$ be the space of continuous piecewise linear finite elements on ω_i. The local contribution of the interpolation operator of Clément is the solution $p_i \in P_1(\omega_i)$ of

$$\int_{\omega_i} (v - p_i)(\boldsymbol{x})q(\boldsymbol{x})\, d\boldsymbol{x} = 0 \quad \forall\, q \in P_1(\omega_i). \tag{C.16}$$

If $v \in L^2(\omega_i)$, then (C.16) is a local $L^2(\omega_i)$ projection. The Clément interpolation operator is defined by

$$P^h_{\text{Cle}} v(\boldsymbol{x}) = \sum_{i=1}^{N} p_i(V_i)\phi_i^h(\boldsymbol{x}), \tag{C.17}$$

where $\{\phi_i^h\}_{i=1}^N$ is the standard basis of the global finite element space P_1. Since $P^h_{\text{Cle}} v(\boldsymbol{x})$ is a linear combination of basis functions of P_1, it defines a map $P^h_{\text{Cle}} : L^1(\Omega) \to P_1$. □

Theorem C.18 (Interpolation Estimate) *Let $k, l \in \mathbb{N} \cup \{0\}$ and $q \in \mathbb{R}$ with $k \le l \le 2$, $1 \le q \le \infty$, and let ω_K be the union of all subdomains ω_i that contain the mesh cell K, see Fig. C.1. Then, it holds for all $v \in W^{l,q}(\omega_K)$ the estimate*

$$\left\| D^k(v - P^h_{\text{Cle}} v) \right\|_{L^q(K)} \le C h^{l-k} \left\| D^l v \right\|_{L^q(\omega_K)}, \tag{C.18}$$

with $h = diam(\omega_K)$, where the constant C is independent of v and h.

Proof The statement of the lemma is obvious in the case $k = l = 2$ since it is $D^2 P^h_{\text{Cle}} v(\boldsymbol{x})|_K = 0$.

Let $k \in \{0, 1\}$. Since $P_1(\omega_K) \subset L^2(\omega_K)$ and because the $L^2(\omega_i)$ projection gives an element with best approximation, one gets with (C.16)

$$P^h_{\text{Cle}} p = p \quad \text{in } K \quad \forall\, p \in P_1(\omega_K). \tag{C.19}$$

Hence, P^h_{Cle} is a consistent operator.

The next step consists in proving the stability of P^h_{Cle}. One obtains with the inverse inequality, see (C.35) below,

$$\|p\|_{L^\infty(\omega_i)} \le Ch^{-d/2}\,\|p\|_{L^2(\omega_i)} \qquad \text{for all } p \in P_1(\omega_i).$$

The inverse inequality and definition (C.16) with the test function $q = p_i$ gives

$$\|p_i\|^2_{L^\infty(\omega_i)} \le Ch^{-d}\,\|p_i\|^2_{L^2(\omega_i)} \le Ch^{-d}\,\|v\|_{L^1(\omega_i)}\,\|p_i\|_{L^\infty(\omega_i)}\,.$$

Dividing by $\|p_i\|_{L^\infty(\omega_i)}$ and applying Hölder's inequality, one obtains for $p^{-1} = 1 - q^{-1}$

$$\begin{aligned}|p_i(V_i)| &\le \|p_i\|_{L^\infty(\omega_i)} \le Ch^{-d}\,\|v\|_{L^1(\omega_i)} = Ch^{-d}\,\|1v\|_{L^1(\omega_i)} \\ &\le Ch^{-d}\,\|v\|_{L^q(\omega_i)}\,\underbrace{\|1\|_{L^p(\omega_i)}}_{=Ch^{d/p}} = Ch^{d(1/p-1)}\,\|v\|_{L^q(\omega_i)} = Ch^{-d/q}\,\|v\|_{L^q(\omega_i)}\end{aligned} \tag{C.20}$$

for all $V_i \in K$. From the regularity of the triangulation, it follows for the basis functions that (inverse estimate)

$$\left\|D^k\phi_i\right\|_{L^\infty(K)} \le Ch^{-k}, \quad k = 0, 1. \tag{C.21}$$

Using the triangle inequality and combining (C.20) and (C.21) yields the stability of P^h_{Cle}

$$\begin{aligned}\left\|D^k P^h_{\text{Cle}} v\right\|_{L^q(K)} &\le \sum_{V_i\in K} |p_i(V_i)|\,\left\|D^k\phi_i\right\|_{L^q(K)} \\ &\le C\sum_{V_i\in K} h^{-d/q}\,\|v\|_{L^q(\omega_i)}\,\left\|D^k\phi_i\right\|_{L^\infty(K)}\,\|1\|_{L^q(K)} \\ &\le C\sum_{V_i\in K} h^{-d/q}\,\|v\|_{L^q(\omega_i)}\,h^{-k}h^{d/q} \\ &= Ch^{-k}\,\|v\|_{L^q(\omega_K)}\,.\end{aligned} \tag{C.22}$$

The remainder of the proof follows the proof of the interpolation error estimate for the polynomial interpolation, Theorem C.8, apart from the fact that a reference cell is not used for the Clément interpolation operator. Using Lemmas C.2 and C.3, one can find a polynomial $p \in P_1(\omega_K)$ with

$$\left\|D^j(v-p)\right\|_{L^q(\omega_K)} \le Ch^{l-j}\,\left\|D^l v\right\|_{L^q(\omega_K)}, \quad 0 \le j \le l \le 2. \tag{C.23}$$

With (C.19), the triangle inequality, $\|\cdot\|_{L^q(K)} \le \|\cdot\|_{L^q(\omega_K)}$, (C.22), and (C.23), one obtains

$$\begin{aligned}\left\|D^k\left(v - P^h_{\text{Cle}}v\right)\right\|_{L^q(K)} &= \left\|D^k\left(v - p + P^h_{\text{Cle}}p - P^h_{\text{Cle}}v\right)\right\|_{L^q(K)} \\ &\le \left\|D^k(v-p)\right\|_{L^q(K)} + \left\|D^k P^h_{\text{Cle}}(v-p)\right\|_{L^q(K)} \\ &\le \left\|D^k(v-p)\right\|_{L^q(\omega_K)} + Ch^{-k}\left\|v-p\right\|_{L^q(\omega_K)} \\ &\le Ch^{l-k}\left\|D^l v\right\|_{L^q(\omega_K)} + Ch^{-k}h^l\left\|D^l v\right\|_{L^q(\omega_K)} \\ &= Ch^{l-k}\left\|D^l v\right\|_{L^q(\omega_K)}.\end{aligned}$$

■

Remark C.19 (Uniform Meshes)

- If all mesh cells in ω_K are of the same size, then h can be replaced by h_K in the interpolation error estimate (C.18).
- If one assumes that the number of mesh cells in ω_K is bounded uniformly for all considered triangulations, the global interpolation estimate

  $$\left\|D^k(v - P^h_{\text{Cle}}v)\right\|_{L^q(\Omega)} \le Ch^{l-k}\left\|D^l v\right\|_{L^q(\Omega)}, \quad 0 \le k \le l \le 2,$$

  follows directly from (C.18).

□

Remark C.20 (Other Finite Element Spaces) The idea of the Clément interpolation can be extended to other finite element spaces, see Clément (1975). In this paper, it is just assumed that the global functionals are values or derivatives of the function in the nodes. Optimal interpolation estimates are given in Clément (1975). □

Remark C.21 (Estimate for Faces) Let $\tilde{\omega}_E$ be the union of all mesh cells who possess at least one vertex which belongs to a face E. Then there holds the interpolation estimate, see Clément (1975) and Verfürth (1996, Lemma 1.4)

$$\left\|v - P^h_{\text{Cle}}v\right\|_{L^2(E)} \le Ch_E^{1/2}\left\|\nabla v\right\|_{L^2(\tilde{\omega}_E)}. \tag{C.24}$$

□

Remark C.22 (Preservation of Homogeneous Dirichlet Boundary Conditions) For global finite element spaces $V^h \subset H^1_0(\Omega)$, it is shown in Clément (1975) that homogeneous Dirichlet boundary conditions can be preserved under some (weak) assumptions on the finite element space. First, the analysis of Clément (1975) is restricted to finite element spaces with certain global functionals as mentioned in Remark C.20. In addition, it is assumed that for the nodes on the boundary the functionals are only values of the function (and no derivatives). For the definition of the global Clément interpolation operator, these values are left unchanged, i.e., equal to zero, and the interpolation is computed for all other degrees of freedom. For this construction, optimal interpolation estimates were proved in Clément (1975).

As a consequence, for finite element spaces $V^h = P_k \cap H_0^1(\Omega)$ or $V^h = Q_k \cap H_0^1(\Omega)$, the Clément interpolant of $v \in H_0^1(\Omega)$ into V^h is well defined and, in particular, the homogeneous Dirichlet boundary values are preserved. □

C.3 Orthogonal Projections

Remark C.23 (On Orthogonal Projections) Let V be a Hilbert space with inner product $(\cdot,\cdot)_V$ and let $V^h \subset V$ be a finite-dimensional space. Another possibility to assign to a function $v \in V$ a function $P^h v \in V^h$ is the orthogonal projection

$$\left(v - P^h v, v^h\right)_V = 0 \quad \forall\ v^h \in V^h.$$

The function $P^h v$ is the best approximation to v with respect to the norm induced by $(\cdot,\cdot)_V$.

This section introduces the most important orthogonal projections together with their properties. A short overview on orthogonal projections can be found in Ern and Guermond (2004, Sect. 1.6.3). □

Definition C.24 ($L^2(\Omega)$ Projection) The $L^2(\Omega)$ projection is defined by

$$P_{L^2}^h \ :\ L^2(\Omega) \to V^h, \quad v \mapsto P_{L^2}^h v$$

with

$$\int_\Omega \left(v - P_{L^2}^h v\right) v^h(\boldsymbol{x})\, d\boldsymbol{x} = \left(v - P_{L^2}^h v, v^h\right) = 0 \quad \forall\ v^h \in V^h. \tag{C.25}$$

□

Definition C.25 (Elliptic Projection, Riesz Projection) The elliptic projection or Riesz projection is defined by

$$P_{H^1}^h \ :\ H^1(\Omega) \to V^h, \quad v \mapsto P_{H^1}^h v$$

with

$$\begin{aligned}&\int_\Omega \left[\left(v - P_{H^1}^h v\right) v^h + \nabla\left(v - P_{H^1}^h v\right)\cdot \nabla v^h\right](\boldsymbol{x})\, d\boldsymbol{x} \\ &= \left(v - P_{H^1}^h v, v^h\right) + \left(\nabla\left(v - P_{H^1}^h v\right), \nabla v^h\right) = 0 \quad \forall\ v^h \in V^h.\end{aligned}$$

□

Lemma C.26 (Stability of the $L^2(\Omega)$ and the Elliptic Projection) *The following stability estimates hold:*

$$\left\|P_{L^2}^h v\right\|_{L^2(\Omega)} \le \|v\|_{L^2(\Omega)} \quad \forall\ v \in L^2(\Omega), \tag{C.26}$$

$$\left\|P_{H^1}^h v\right\|_{L^2(\Omega)} \le \|v\|_{H^1(\Omega)} \quad \forall\ v \in H^1(\Omega). \tag{C.27}$$

Proof Choosing $v^h = P_{L^2}^h v$ in (C.25) and applying the Cauchy–Schwarz inequality (A.10) yields

$$\left\|P_{L^2}^h v\right\|_{L^2(\Omega)}^2 = \left(P_{L^2}^h v, P_{L^2}^h v\right) = \left(v, P_{L^2}^h v\right) \le \|v\|_{L^2(\Omega)} \left\|P_{L^2}^h v\right\|_{L^2(\Omega)}.$$

Dividing by $\left\|P_{L^2}^h v\right\|_{L^2(\Omega)} \neq 0$ gives (C.26). For $\left\|P_{L^2}^h v\right\|_{L^2(\Omega)} = 0$, the validity of (C.26) is obvious.

Estimate (C.27) is proved in the same way as estimate (C.26). ■

Theorem C.27 (Error Estimates for the Orthogonal Projections, Ern and Guermond (2004, Propositions 1.134 and 1.135)) *Consider finite element spaces $V^h = P_k$ or $V^h = Q_k$. Let $k \ge 0$ and $0 \le l \le k$, then there is a constant C, independent of h, such that*

$$\left\|v - P_{L^2}^h v\right\|_{L^2(\Omega)} \le C h^{l+1} |v|_{H^{l+1}(\Omega)} \quad \forall\ v \in H^{l+1}(\Omega),\ \forall\ h. \tag{C.28}$$

If the family of triangulations $\{\mathcal{T}^h\}$ is quasi-uniform, then it is for $k \ge 1$, $1 \le l \le k$

$$\left\|v - P_{L^2}^h v\right\|_{H^1(\Omega)} \le C h^{l} |v|_{H^{l+1}(\Omega)} \quad \forall\ v \in H^{l+1}(\Omega),\ \forall\ h, \tag{C.29}$$

with C independent of h.

If $k \ge 1$ and $1 \le l \le k$, then the estimate

$$\left\|v - P_{H^1}^h v\right\|_{H^1(\Omega)} \le C h^{l} |v|_{H^{l+1}(\Omega)} \quad \forall\ v \in H^{l+1}(\Omega),\ \forall\ h, \tag{C.30}$$

holds with C independent of h. If Ω is convex, then (C.30) *can be improved to*

$$\left\|v - P_{H^1}^h v\right\|_{H^1(\Omega)} \le C h^{l+1} |v|_{H^{l+1}(\Omega)} \quad \forall\ v \in H^{l+1}(\Omega),\ \forall\ h, \tag{C.31}$$

with C independent of h.

Remark C.28 (Jumps Across Faces for the $L^2(\Omega)$ Projection on Piecewise Constant Functions) Let K be a mesh cell and $E \subset \partial K$ be one of its faces. Transforming the integral on E to a face of the reference mesh cell, applying the trace theorem, Theorem A.34, and transforming the integral from the reference mesh cell to K, which is the same procedure as in the proof of Corollary C.15, yields

$$\left\| v - P_{L^2}^h v \right\|_{L^2(E)}^2 \le C \left(h_K^{-1} \left\| v - P_{L^2}^h v \right\|_{L^2(K)}^2 + h_K \left\| \nabla \left(v - P_{L^2}^h v \right) \right\|_{L^2(K)}^2 \right). \tag{C.32}$$

With Young's inequality (A.5), it is

$$\begin{aligned} &\left\| [|v - P_{L^2}^h v|]_E \right\|_{L^2(E)}^2 \\ &= \int_E \left(\left(v - P_{L^2}^h v \right)\big|_{K_1} - \left(v - P_{L^2}^h v \right)\big|_{K_2} \right)^2 \, ds \qquad \text{(C.33)} \\ &\le 2 \int_E \left(\left(v - P_{L^2}^h v \right)\big|_{K_1} \right)^2 \, ds + 2 \int_E \left(\left(v - P_{L^2}^h v \right)\big|_{K_2} \right)^2 \, ds, \end{aligned}$$

where K_1 and K_2 are the mesh cells with the common face E. For quasi-uniform triangulations, one gets from (C.32), (C.33), and (C.28) for $P_{L^2}^h v \in P_0$ (or $P_{L^2}^h v \in Q_0$) and $v \in H^1(\Omega)$

$$\begin{aligned} &\sum_{E \in \mathcal{E}^h} \left\| [|v - P_{L^2}^h v|]_E \right\|_{L^2(E)}^2 \\ &\le C \left(h^{-1} \sum_{E \in \mathcal{E}^h} \left\| v - P_{L^2}^h v \right\|_{L^2(\omega_E)}^2 + h \sum_{E \in \mathcal{E}^h} \sum_{K \in \omega_E} \left\| \nabla \left(v - P_{L^2}^h v \right) \right\|_{L^2(K)}^2 \right) \\ &\le C \left(h^{-1} \left\| v - P_{L^2}^h v \right\|_{L^2(\Omega)}^2 + h \left\| \nabla v \right\|_{L^2(\Omega)}^2 \right) \\ &\le C h \left\| \nabla v \right\|_{L^2(\Omega)}^2 . \qquad \text{(C.34)} \end{aligned}$$

□

C.4 Inverse Estimate

Remark C.29 (On Inverse Estimates) The approach for proving interpolation error estimates can be uses also to prove so-called inverse estimates. With inverse estimates, a norm of a higher order derivative of a finite element function is estimated by a norm of a lower order derivative of this function. Likewise, norms in different Lebesgue spaces are estimated. One obtains as penalty a factor with negative powers of the diameter of the mesh cell. □

Theorem C.30 (Inverse Estimate) *Let* $0 \le k \le l$ *be natural numbers and let* $p, q \in [1, \infty]$. *Then there is a constant* C_{inv}, *which depends only on*

$k, l, p, q, \hat{K}, \hat{P}(\hat{K})$, such that

$$\left\| D^l v^h \right\|_{L^q(K)} \le C_{\text{inv}} h_K^{(k-l)-d(p^{-1}-q^{-1})} \left\| D^k v^h \right\|_{L^p(K)} \quad \forall\ v^h \in P(K). \tag{C.35}$$

Proof In the first step, (C.35) is shown for $h_{\hat{K}} = 1$ and $k = 0$ on the reference mesh cell. Since all norms are equivalent in finite-dimensional spaces, one obtains

$$\left\| D^l \hat{v}^h \right\|_{L^q(\hat{K})} \le \left\| \hat{v}^h \right\|_{W^{l,q}(\hat{K})} \le C \left\| \hat{v}^h \right\|_{L^p(\hat{K})} \quad \forall\ \hat{v}^h \in \hat{P}(\hat{K}). \tag{C.36}$$

If $k > 0$, then one sets

$$\tilde{P}(\hat{K}) = \left\{ \partial_{\boldsymbol{\alpha}} \hat{v}^h \ :\ \hat{v}^h \in \hat{P}(\hat{K}), |\boldsymbol{\alpha}| = k \right\},$$

which is also a space consisting of polynomials. The application of (C.36) to $\tilde{P}(\hat{K})$ gives

$$\left\| D^l \hat{v}^h \right\|_{L^q(\hat{K})} = \sum_{|\boldsymbol{\alpha}|=k} \left\| D^{l-k} \left(\partial_{\boldsymbol{\alpha}} \hat{v}^h \right) \right\|_{L^q(\hat{K})} \le C \sum_{|\boldsymbol{\alpha}|=k} \left\| \partial_{\boldsymbol{\alpha}} \hat{v}^h \right\|_{L^p(\hat{K})} = C \left\| D^k \hat{v}^h \right\|_{L^p(\hat{K})}.$$

This estimate is transformed to an arbitrary mesh cell K analogously as for the interpolation error estimates, compare the proof of Theorem C.13. From the estimates for the transformations, one obtains

$$\begin{aligned} \left\| D^l v^h \right\|_{L^q(K)} &\le C h_K^{-l+d/q} \left\| D^l \hat{v}^h \right\|_{L^q(\hat{K})} \le C h_K^{-l+d/q} \left\| D^k \hat{v}^h \right\|_{L^p(\hat{K})} \\ &\le C_{\text{inv}} h_K^{k-l+d/q-d/p} \left\| D^k v^h \right\|_{L^p(K)}. \end{aligned}$$

■

Remark C.31 (On the Proof) The crucial point in the proof is the equivalence of all norms in finite-dimensional spaces. Such a property does not hold in infinite-dimensional spaces. □

Corollary C.32 (Global Inverse Estimate) *Let $p = q$ and let $\{\mathcal{T}^h\}$ be a quasi-uniform family of triangulations of Ω, then*

$$\left\| D^l v^h \right\|_{L^{p,h}(\Omega)} \le C_{\text{inv}} h^{k-l} \left\| D^k v^h \right\|_{L^{p,h}(\Omega)}, \tag{C.37}$$

where

$$\|\cdot\|_{L^{p,h}(\Omega)} = \left(\sum_{K \in \mathcal{T}^h} \|\cdot\|_{L^p(K)}^p \right)^{1/p}.$$

Remark C.33 (On $\|\cdot\|_{L^{p,h}(\Omega)}$) The cell-wise definition of the norm is important for $k \geq 2$ or $l \geq 2$ since in these cases finite element functions generally do not possess the regularity for the global norm to be well defined. It is also important for $l \geq 1$ and non-conforming finite element functions. □

Appendix D
Examples for Numerical Simulations

Remark D.1 (General Considerations) The definition of good test examples is of importance for the assessment of numerical schemes. There are different classes of test problems:

- *Academic test examples with prescribed solution.* In these examples, the velocity field $\boldsymbol{u}$ and the pressure p are prescribed by analytical functions. The right-hand side, boundary conditions, and the initial condition are chosen such that the strong form of the considered equation is fulfilled.
  These examples serve for supporting convergence estimates. A connection to real life problems is generally not given.
  In the definition of these examples, one has to take care that the velocity field is divergence-free. Depending on the type of boundary condition, see Sect. 2.4, the integral mean value of the pressure has to vanish or the integral of the Dirichlet boundary data has to satisfy the compatibility condition (2.33).
- *Academic test examples with features of real life flows.* These examples contain on the one hand some important features of real life flow problems, but on the other hand, a number of simplifications are used to facilitate their implementation and the assessment of the results. Generally, an analytical solution is not known. Reference values for quantities of interest are obtained by performing simulations on very fine grids in space and time.
- *Real life examples.* A proposed numerical method should work well for this type of examples. However, often the data are incomplete in real life examples, e.g., temporally and spatially resolved boundary conditions are generally not known. In addition, reference values to compare with are coming often from measurements. These values are generally mean values in space or in time. A certain measurement error has always to be expected.
  In conclusion, special care and special techniques are necessary to assess numerical methods at real world problems. Generally, none of the methods will produce results that agree completely with the real world flow, already because of

V. John, *Finite Element Methods for Incompressible Flow Problems*, Springer Series in Computational Mathematics 51, DOI 10.1007/978-3-319-45750-5

the incomplete data. Thus, one has to evaluated the differences in an appropriate sense.

Appendix D presents some examples that were used in the numerical illustrations presented in this monograph. Most of them are standard test problems from the literature. The simulations were performed with the code MooNMD, see John and Matthies (2004). □

Remark D.2 (Forces Exerted by the Flow on Bodies) Consider a body with boundary Γ_{body} in a flow. Then, the force that is exerted by the flow on the body in the (unit) direction $\boldsymbol{w}$ is given by

$$F_{\boldsymbol{w}} = -\int_{\Gamma_{\text{body}}} (\mathbb{S}\boldsymbol{n}) \cdot \boldsymbol{w}\, ds = -\int_{\Gamma_{\text{body}}} ((2\mu\mathbb{D}(\boldsymbol{v}) - P\mathbb{I})\,\boldsymbol{n}) \cdot \boldsymbol{w}\, ds \quad [\text{N}]\,.$$

Here, $\boldsymbol{n}$ is the unit normal vector that points outward with respect to the flow domain Ω and the stress tensor is given in (2.18). Thus, the force is defined by the projection of the (negative of the) normal stress into the direction $\boldsymbol{w}$. For the dimensionless quantities, one obtains with (2.23)

$$\begin{aligned} F_{\boldsymbol{w}} &= -L^{d-1}\int_{\Gamma_{\text{body}}} \left(\left(\frac{2\mu U}{L}\mathbb{D}(\boldsymbol{u}) - \rho U^2 p\mathbb{I}\right)\boldsymbol{n}\right) \cdot \boldsymbol{w}\, ds \\ &= -\rho U^2 L^{d-1}\int_{\Gamma_{\text{body}}} \left(\left(\frac{2\mu}{\rho U L}\mathbb{D}(\boldsymbol{u}) - p\mathbb{I}\right)\boldsymbol{n}\right) \cdot \boldsymbol{w}\, ds \\ &= -\rho U^2 L^{d-1}\int_{\Gamma_{\text{body}}} ((2\nu\mathbb{D}(\boldsymbol{u}) - p\mathbb{I})\,\boldsymbol{n}) \cdot \boldsymbol{w}\, ds, \end{aligned} \qquad \text{(D.1)}$$

where $\nu = \text{Re}^{-1}$. The factor $1/L$ in the first line comes from the transform of the spatial derivatives and the factor L^{d-1} from the transform of the integral.

Formula (D.1) fits well into a variational formulation of the Navier–Stokes equations. Assuming that Γ_{body} is not attached to the rest of the boundary, $\overline{\Gamma_{\text{body}}} \cap \overline{\Gamma \setminus \Gamma_{\text{body}}} = \emptyset$, one can extend $\boldsymbol{w}$ to a function in Ω such that $\boldsymbol{w} \in H^1(\Omega)$ and $\boldsymbol{w}$ vanishes at $\overline{\Gamma \setminus \Gamma_{\text{body}}}$. Testing the momentum equation of (2.25) with $\boldsymbol{w}$ and applying integration by parts gives

$$\begin{aligned} &(\partial_t \boldsymbol{u}, \boldsymbol{w}) + (2\nu\mathbb{D}(\boldsymbol{u}), \mathbb{D}(\boldsymbol{w})) + ((\boldsymbol{u}\cdot\nabla)\boldsymbol{u}, \boldsymbol{w}) - (\nabla\cdot\boldsymbol{w}, p) \\ &\quad -\langle \boldsymbol{f}, \boldsymbol{w}\rangle_{(H^1(\Omega))', H^1(\Omega)} = \int_{\Gamma_{\text{body}}} ((2\nu\mathbb{D}(\boldsymbol{u}) - p\mathbb{I})\,\boldsymbol{n}) \cdot \boldsymbol{w}\, ds, \end{aligned}$$

such that

$$F_{\boldsymbol{w}} = -\rho U^2 L^{d-1}\Big[(\partial_t \boldsymbol{u}, \boldsymbol{w}) + (2\nu \mathbb{D}(\boldsymbol{u}), \mathbb{D}(\boldsymbol{w})) + ((\boldsymbol{u}\cdot\nabla)\boldsymbol{u}, \boldsymbol{w}) \\ - (\nabla\cdot \boldsymbol{w}, p) - \langle \boldsymbol{f}, \boldsymbol{w}\rangle_{(H^1(\Omega))', H^1(\Omega)}\Big]. \quad \text{(D.2)}$$

The term with the time derivative vanishes for steady-state problems. For the viscous and convective term there are different forms, see (4.5) and Sect. 6.1.2. In the continuous setting, these forms are equivalent, since $\boldsymbol{u}$ is weakly divergence-free, and they lead to equivalent formulas. However, the equivalence is usually lost for finite element methods such that the use of different forms for the viscous and convective term leads to different values for the discrete approximation of $F_{\boldsymbol{w}}$. To the best of our knowledge, there are no investigations on the size of these differences.

In practice, $F_{\boldsymbol{w}}$ is computed with (D.2) also if the body is attached to the rest of the boundary of the flow domain, e.g., for the flows around a cylinder in John (2002), see also Example D.6, and around a wall-mounted cube (Example D.14) in Hoffman (2005), Hoffman and Johnson (2006).

Of importance in applications is the force in main flow direction, the so-called drag force. Based on this force, a dimensionless coefficient, the drag coefficient is defined by

$$c_{\text{drag}} = \frac{2F_{\text{drag}}}{\rho A V^2}, \quad \text{(D.3)}$$

where V is a measure for the speed of the obstacle relative to the fluid and A is a reference area. The quantity $\rho V^2/2$ is called dynamic pressure. With (D.1), one obtains

$$c_{\text{drag}} = -\frac{2U^2 L^{d-1}}{AV^2} \int_{\Gamma_{\text{body}}} ((2\nu\mathbb{D}(\boldsymbol{u}) - p\mathbb{I})\boldsymbol{n}) \cdot \boldsymbol{w}_{\text{drag}}\, ds \quad \text{(D.4)}$$

and with (D.2), one gets

$$c_{\text{drag}} = -\frac{2U^2 L^{d-1}}{AV^2}\Big[\big(\partial_t \boldsymbol{u}, \boldsymbol{w}_{\text{drag}}\big) + \big(2\nu \mathbb{D}(\boldsymbol{u}), \mathbb{D}\big(\boldsymbol{w}_{\text{drag}}\big)\big) + \big((\boldsymbol{u}\cdot\nabla)\boldsymbol{u}, \boldsymbol{w}_{\text{drag}}\big) \\ - \big(\nabla\cdot \boldsymbol{w}_{\text{drag}}, p\big) - \langle \boldsymbol{f}, \boldsymbol{w}_{\text{drag}}\rangle_{(H^1(\Omega))', H^1(\Omega)}\Big]. \quad \text{(D.5)}$$

Similarly important is the force exerted perpendicular to the main flow direction, the so-called lift force. Let the main flow direction be given by $\boldsymbol{w}_{\text{drag}} = (w_1, w_2, 0)^T$, then the orthogonal direction is set to be $\boldsymbol{w}_{\text{lift}} = (-w_2, w_1, 0)^T$. The dimensionless lift coefficient is defined by

$$c_{\text{lift}} = \frac{2F_{\text{lift}}}{\rho A V^2}, \quad \text{(D.6)}$$

such that one gets the representations

$$c_{\text{lift}} = -\frac{2U^2L^{d-1}}{AV^2}\int_{\Gamma_{\text{body}}} ((2\nu\mathbb{D}(\boldsymbol{u}) - p\mathbb{I})\,\boldsymbol{n}) \cdot \boldsymbol{w}_{\text{lift}}\, ds \tag{D.7}$$

with (D.1) and

$$\begin{aligned} c_{\text{lift}} = -\frac{2U^2L^{d-1}}{AV^2}\Big[&(\partial_t\boldsymbol{u}, \boldsymbol{w}_{\text{lift}}) + (2\nu\mathbb{D}(\boldsymbol{u}), \mathbb{D}(\boldsymbol{w}_{\text{lift}})) + ((\boldsymbol{u}\cdot\nabla)\boldsymbol{u}, \boldsymbol{w}_{\text{lift}}) \\ &- (\nabla\cdot\boldsymbol{w}_{\text{lift}}, p) - \langle \boldsymbol{f}, \boldsymbol{w}_{\text{lift}}\rangle_{(H^1(\Omega))', H^1(\Omega)}\Big] \end{aligned} \tag{D.8}$$

with (D.2).

In the literature, one can find also other formulas for the drag and lift force, e.g., in Schäfer and Turek (1996) formulas are given for the special case $\boldsymbol{w}_{\text{drag}} = (1, 0, 0)^T$ and $\partial_z u_3|_{\Gamma_{\text{body}}} = 0$. □

D.1 Examples for Steady-State Flow Problems

Example D.3 (A Two-dimensional Steady-State Example with Prescribed Solution in the Unit Square and with Homogeneous Dirichlet Boundary Conditions) Consider $\Omega = (0, 1)^2$ and the stream function

$$\phi = 1000x^2(1-x)^4y^3(1-y)^2.$$

Then, the velocity field is defined by

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} \partial_y\phi \\ -\partial_x\phi \end{pmatrix} = 1000\begin{pmatrix} x^2(1-x)^4y^2(1-y)(3-5y) \\ -2x(1-x)^3(1-3x)y^3(1-y)^2 \end{pmatrix}. \tag{D.9}$$

It follows, using the Theorem of Schwarz, that

$$\nabla\cdot\boldsymbol{u} = \partial_x u_1 + \partial_y u_2 = \partial_{xy}\phi - \partial_{yx}\phi = \partial_{xy}\phi - \partial_{xy}\phi = 0.$$

The equations will be equipped with Dirichlet boundary conditions on the whole boundary. It is $\boldsymbol{u} = \boldsymbol{0}$ on Γ such that the compatibility condition (2.33) is fulfilled.

Because of the Dirichlet boundary conditions on Γ, the pressure should be in $L_0^2(\Omega)$. This is the only essential requirement. The following pressure was chosen for the definition of this example

$$p = \pi^2(xy^3\cos(2\pi x^2y) - x^2y\sin(2\pi xy)) + \frac{1}{8}.$$

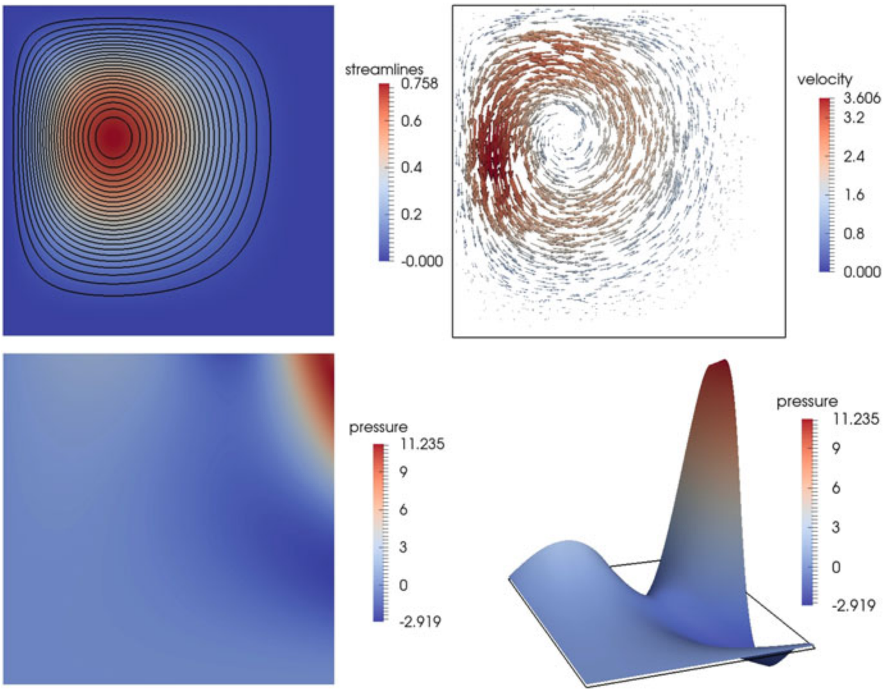

Fig. D.1 Example D.3. Stream function (*top left*) velocity (*top right*) and pressure (*bottom*). These plots are based on results obtained with numerical simulations

The stream function, velocity, and pressure are presented in Fig. D.1. It can be seen that the velocity field consists essentially of one big vortex.

Because of the homogeneous Dirichlet boundary conditions on $\varGamma$, this example fits into the framework of many results obtained in the numerical analysis of finite element methods. □

Example D.4 (The Driven Cavity Problem in Two Dimensions) The two-dimensional driven cavity problem is probably the most popular test example for two-dimensional steady-state flows. It is defined in $\varOmega = (0,1)^2$. There are no body forces, i.e., the right-hand side of the Navier–Stokes equations (6.1) is $\boldsymbol{f} = \boldsymbol{0}$, the boundary conditions of the classical driven cavity problem are prescribed in the literature often in the form

$$\boldsymbol{u} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{at } x = 0, x = 1, y = 0; \quad \boldsymbol{u} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{at } y = 1, \tag{D.10}$$

and the concrete examples are defined with the Reynolds number $\mathrm{Re} = \nu^{-1}$.

The boundary condition (D.10) has a jump at the upper corners of the domain. If pairs of finite element spaces like the Taylor–Hood pair are used, then one needs the specification of the values of $\boldsymbol{u}$ in these corners. Such a specification can be found only in some papers, e.g., in Olshanskii (2002) the values $\boldsymbol{u}(0,1) = \boldsymbol{u}(1,1) = (1,0)^T$ were used. Independently of the concrete choice of $\boldsymbol{u}(0,1)$ and $\boldsymbol{u}(1,1)$, the Dirichlet boundary condition does not belong to $H^{1/2}(\Gamma)$ and by the trace theorem, Theorem A.34 for $s = 1$, the velocity cannot belong to $H^1(\Omega)$. Consequently, the classical driven cavity problem does not fit into the framework of Chap. 6. Strictly speaking, the solution is not sufficiently regular for the application of conforming finite element methods.

Concerning reference solutions for the classical driven cavity problem, it is usually referred to Ghia et al. (1982). Streamlines, profiles of the solution on cut lines, and the position of vortex centers are provided for $\mathrm{Re} \in \{100, 400, 1000, 3200, 5000, 7500, 10{,}000\}$. However, it was shown, e.g., in Cazemier et al. (1998), Tiesinga et al. (2002), Bruneau and Saad (2006), that in the classical lid driven cavity example, the stationary solution becomes unstable at a Reynolds number of around $\mathrm{Re} = 8000$ and that at $\mathrm{Re} = 10000$ there is a stable periodic solution. Thus, numerical methods that compute a stable steady-state solution for $\mathrm{Re} = 10000$ have to introduce additional numerical viscosity. In fact, the simulations in Ghia et al. (1982) were performed with a finite difference scheme and some stabilization of upwind type.

To avoid the irregularity of the solution at the upper corners, regularized driven cavity problems have been proposed, e.g., in Elman et al. (2014, p. 125). In the definition of such regularized problems, one has to pay attention that $\nabla \cdot \boldsymbol{u} = 0$ holds also on the boundary and that the divergence at the corners is completely defined by the prescribed boundary condition. A regularized driven cavity problem that satisfies this requirement was proposed in de Frutos et al. (2016c), which is equipped with the boundary condition

$$\boldsymbol{u}(x,1) = \begin{pmatrix} u_1(x) \\ 0 \end{pmatrix}, \tag{D.11}$$

$$u_1(x) = \begin{cases} 1 - \frac{1}{4}\left(1 - \cos\left(\frac{x_1 - x}{x_1}\pi\right)\right)^2 & \text{for } x \in [0, x_1], \\ 1 & \text{for } x \in (x_1, 1 - x_1), \\ 1 - \frac{1}{4}\left(1 - \cos\left(\frac{x - (1 - x_1)}{x_1}\pi\right)\right)^2 & \text{for } x \in [1 - x_1, 1]. \end{cases}$$

All simulations of this example presented in this monograph were performed with $x_1 = 0.1$. The solutions of this problem for $\mathrm{Re} = 1$ and $\mathrm{Re} = 3200$ are presented in Fig. D.2. In the low Reynolds number case $\mathrm{Re} = 1$, there is a big vortex whose center is close to the upper boundary. The pressure shows large peaks at both upper corners. For $\mathrm{Re} = 3200$, the velocity field consists of one big vortex whose center is located near the center of the cavity. There are smaller counter-rotating vortices in both lower corners and at the upper part of the left boundary. The pressure has a positive peak in the upper right corner. □

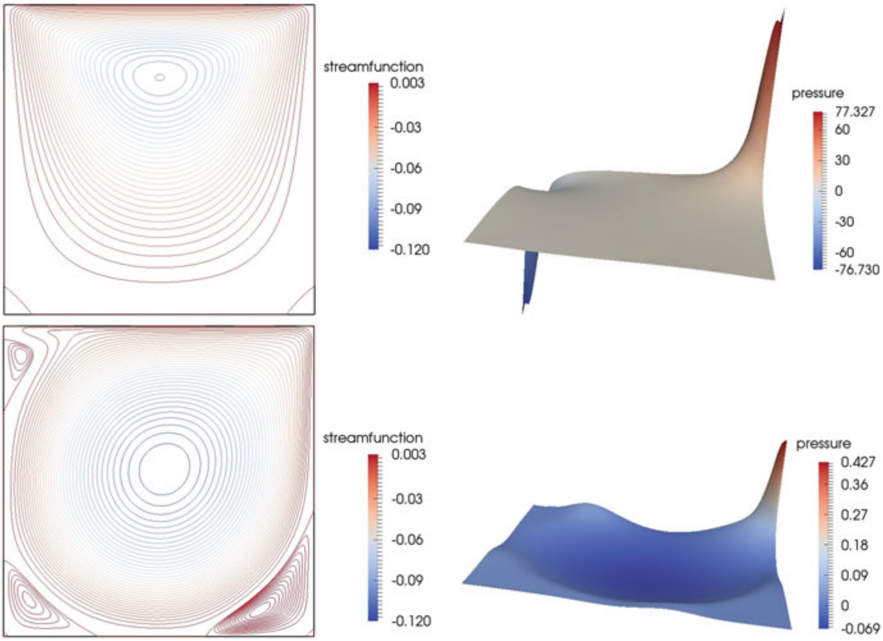

Fig. D.2 Example D.4, regularized driven cavity problem with boundary conditions (D.11). Stream function (*left*, 36 intervals in $[-0.12, 0]$, 15 intervals in $[0, 0.03]$) and pressure (*right*) for $\mathrm{Re} = 1$ (*top*) and $\mathrm{Re} = 3200$ (*bottom*)

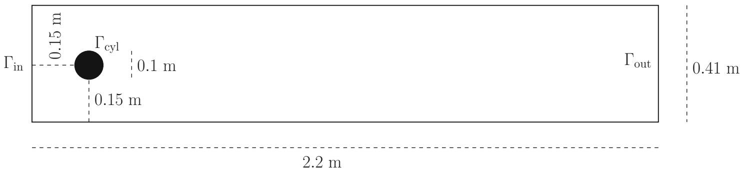

Fig. D.3 Example D.5. Domain

Example D.5 (A Steady-State Flow Around a Cylinder at Re = 20 in Two Dimensions) This example was defined in Schäfer and Turek (1996). It considers a flow in a two-dimensional domain with a two-dimensional cylinder (circle), see Fig. D.3 for a sketch of this domain.

The dynamic viscosity of the fluid is given by $\mu = 10^{-3}\,\mathrm{Pa\,s}$ and its density by $\rho = 1\,\mathrm{kg/m^3}$. These values are approximately the coefficients for water. The parabolic inflow profile is defined by

$$\boldsymbol{v}(0\,\mathrm{m}, y) = \frac{1}{0.41^2}\begin{pmatrix} 1.2y(0.41 - y) \\ 0 \end{pmatrix}\,\mathrm{m/s}, \quad 0\,\mathrm{m} \le y \le 0.41\,\mathrm{m}. \tag{D.12}$$

At the top and the bottom of the channel and at the surface Γ_{body} of the cylinder, no-slip boundary conditions are prescribed. With respect to the outlet there are two possible boundary conditions in this example. Either, a parabolic outflow boundary condition

$$\boldsymbol{v}(2.2\,\text{m}, y) = \frac{1}{0.41^2}\begin{pmatrix}1.2y(0.41-y)\\ 0\end{pmatrix}\,{}^{\text{m}}/_{\text{s}}, \quad 0\,\text{m} \le y \le 0.41\,\text{m},$$

is applied or the do-nothing boundary condition

$$\mathbb{S}\boldsymbol{n} = (-\mu\nabla\boldsymbol{v} + P\mathbb{I})\,\boldsymbol{n} = \boldsymbol{0}\,{}^{\text{N}}/_{\text{m}^2} \quad \text{on } \Gamma_{\text{out}} \tag{D.13}$$

is used, where $\boldsymbol{n}$ is the outward pointing unit normal vector. The mean inflow velocity is given by

$$U_{\text{mean}} = \frac{1}{0.41^2}\,\frac{\int_0^{0.41} 1.2y(0.41-y)\,dy}{\int_0^{0.41}\,dy}\,{}^{\text{m}}/_{\text{s}} = \frac{1}{5}\frac{0.41^3}{0.41^3}\,{}^{\text{m}}/_{\text{s}} = 0.2\,{}^{\text{m}}/_{\text{s}}.$$

Based on the mean inflow, the diameter $d = 0.1$ m of the cylinder, and the kinematic viscosity μ/ρ, the Reynolds number of the flow is $\text{Re} = 20$. There are no external forces acting on the flow, i.e., $\boldsymbol{f}_{\text{ext}} = \boldsymbol{0}\,{}^{\text{N}}/_{\text{m}^3}$. This setup admits a stable steady-state solution, see Figs. D.4 and D.5.

Using the characteristic length scale $L = 1$ m and the characteristic velocity scale $U = 1\,{}^{\text{m}}/_{\text{s}}$, one obtains the steady-state Navier–Stokes equations (6.1) without dimensions with $\nu = \mu/(\rho U L) = 10^{-3}$, the inflow condition

$$\boldsymbol{u}(0, y) = 0.41^{-2}\left(\begin{pmatrix}1.2y(0.41-y)\\ 0\end{pmatrix}\right), \quad 0 \le y \le 0.41,$$

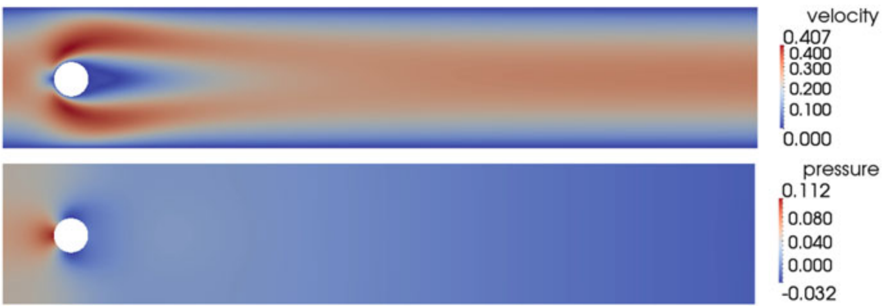

Fig. D.4 Example D.5. Absolute value of the velocity (*top*) and pressure (*bottom*) in the flow around a cylinder

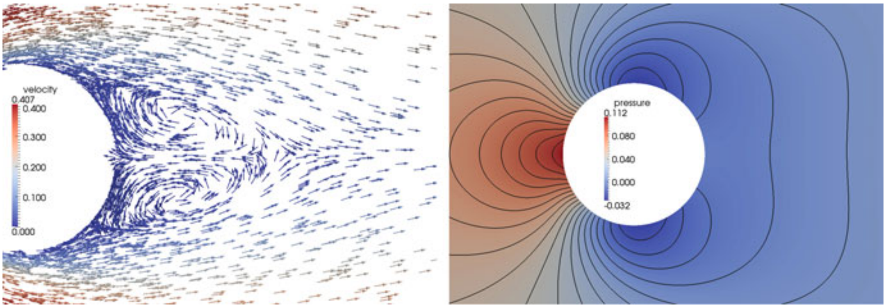

Fig. D.5 Example D.5. Velocity and pressure at the cylinder

and the outflow condition

$$\boldsymbol{u}(2.2, y) = 0.41^{-2}\left(\begin{pmatrix} 1.2y(0.41-y) \\ 0 \end{pmatrix}\right), \quad 0 \le y \le 0.41, \tag{D.14}$$

or

$$\mathbb{S}\boldsymbol{n} = (-\nu\nabla\boldsymbol{u} + p\mathbb{I})\,\boldsymbol{n} = \boldsymbol{0} \quad \text{on } \Gamma_{\text{out}}. \tag{D.15}$$

Quantities of interest in this example are the forces that are exerted on the body, concretely, the drag coefficient given in (D.3) and the lift coefficient given in (D.6). For the considered two-dimensional example, the main flow direction is $\boldsymbol{w}_{\text{drag}} = (1,0)^T$, the orthogonal direction is $\boldsymbol{w}_{\text{lift}} = (0,1)^T$, the relative speed is taken as $V = U_{\text{mean}}$, and the reference area is the diameter of the cylinder $A = d$.

In Remark D.2, two ways are described for evaluating the drag and lift coefficient. In this example, Γ_{body} is a circle, which cannot be represented exactly in finite element methods. Hence, the commitment of a substantial error can be expected if the boundary integral formulations (D.4) and (D.7) are used, just arising from the approximation of the boundary. Thus, the volume (area in two dimensions) integral formulations (D.5) and (D.8) should be preferred. Since the considered example is stationary and $\boldsymbol{f} = \boldsymbol{0}$, the drag coefficient can be computed with

$$c_{\text{drag}} = -500\big((\nu\nabla\boldsymbol{u}, \nabla\boldsymbol{w}_{\text{drag}}) + n(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{w}_{\text{drag}}) - (\nabla\cdot\boldsymbol{w}_{\text{drag}}, p)\big) \tag{D.16}$$

for any function $\boldsymbol{w}_{\text{drag}} \in H^1(\Omega)$ with $\boldsymbol{w}_{\text{drag}} = \boldsymbol{0}$ on $\Gamma \setminus \Gamma_{\text{body}}$ and $\boldsymbol{w}_{\text{drag}}|_{\Gamma_{\text{body}}} = (1,0)^T$. In (D.16), the typical form of the viscous term utilized in steady-state simulations is used and $n(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{w}_{\text{drag}})$ can be any of the forms of the convective term from Sect. 6.1.2. Similarly, the lift coefficient can be computed by

$$c_{\text{lift}} = -500\big((\nu\nabla\boldsymbol{u}, \nabla\boldsymbol{w}_{\text{lift}}) + n(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{w}_{\text{lift}}) - (\nabla\cdot\boldsymbol{w}_{\text{lift}}, p)\big) \tag{D.17}$$

for any function $\boldsymbol{w}_{\text{lift}} \in H^1(\Omega)$ with $\boldsymbol{w}_{\text{lift}} = \boldsymbol{0}$ on $\Gamma \setminus \Gamma_{\text{body}}$ and $\boldsymbol{w}_{\text{lift}}|_{\Gamma_{\text{body}}} = (0,1)^T$.

Another quantity of interest is the difference of the pressure between the front and the back of the cylinder

$$\Delta P = P(0.15, 0.2) - P(0.25, 0.2),$$

which becomes for the dimensionless variables

$$\Delta P = \rho U^2 \big(p(0.15, 0.2) - p(0.25, 0.2) \big). \tag{D.18}$$

Rather accurate reference values are known for the coefficients in the case of Dirichlet boundary conditions at the outlet (D.14) from John and Matthies (2001)

$$c_{\text{drag,ref}} = 5.57953523384, \tag{D.19}$$

$$c_{\text{lift,ref}} = 0.010618937712, \tag{D.20}$$

$$\Delta p_{\text{ref}} = 0.11752016697. \tag{D.21}$$

Using the do-nothing outflow boundary condition (D.15), the reference values for the drag coefficient and the pressure difference are the same. Only the lift coefficient is a little bit sensitive to the boundary condition at Γ_{out}. A reference value in this case was found in Nabh (1998) to be

$$c_{\text{lift,ref}} = 0.010618948146. \tag{D.22}$$

□

Example D.6 (A Steady-State Flow Around a Cylinder at Re = 20 in Three Dimensions) This problem is an extension of Example D.5 to three dimensions. It was also defined in Schäfer and Turek (1996). The domain Ω is the channel with a cylinder shown in Fig. D.6. The height of the channel is $H = 0.41$ m and the diameter of the cylinder $d = 0.1$ m. There are no external forces such that the right-hand side of the momentum equation vanishes, i.e., $\boldsymbol{f} = \boldsymbol{0}$. The dynamic viscosity of the fluid is given by $\mu = 10^{-3}$ Pa s and its density by $\rho = 1\,{}^{\text{kg}}/_{\text{m}^3}$. The inflow condition is prescribed by

$$\boldsymbol{v}|_{\text{in}} = \frac{1}{H^4} \begin{pmatrix} 16Uyz(H-y)(H-z) \\ 0 \\ 0 \end{pmatrix} {}^{\text{m}}/_{\text{s}},$$

with $U = 0.45\,{}^{\text{m}}/_{\text{s}}$. Based on ν, d, and the mean inflow

$$U_{\text{mean}} = \frac{\int_0^H \int_0^H \boldsymbol{v}|_{\text{in}}\, dydz}{\int_0^H \int_0^H dydz} = \frac{16 \cdot 0.45}{36} = 0.2\,{}^{\text{m}}/_{\text{s}},$$

the flow has the Reynolds number Re = 20.

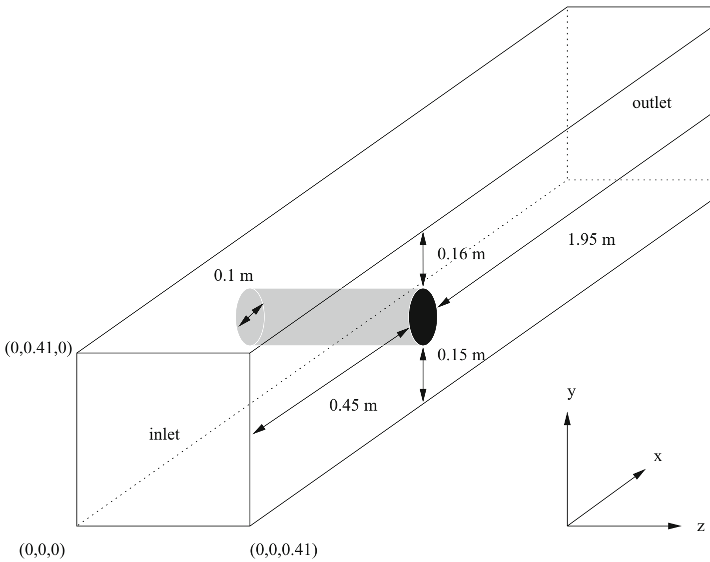

Fig. D.6 Example D.6. Domain

With the characteristic length scale $L = 1\,\mathrm{m}$ and the characteristic velocity scale $U = 1\,\mathrm{m/s}$, the non-dimensional steady-state Navier–Stokes equations (6.1) are derived with $\nu = \mu/(\rho UL) = 10^{-3}$. At the outlet, either the same boundary condition as at the inlet can be used or some form of the do-nothing condition. At all other boundaries, no-slip conditions $\boldsymbol{u} = \boldsymbol{0}$ are prescribed. As in the two-dimensional situation, Example D.5, counter-rotating vortices form behind the cylinder. Considering a cut plane, e.g., $z = 0.41/2$, the flow field looks similarly to the field depicted in Figs. D.4 and D.5.

Quantities of interest are the drag and the lift coefficient at the cylinder and the pressure difference

$$\Delta p = p(0.45, 0.2, 0.205) - p(0.55, 0.2, 0.205). \tag{D.23}$$

The computation of drag and lift coefficients was discussed in detail in Remark D.2. Because finite element methods need to approximate the cylinder, the use of volume integrals should be preferred. However, in this example there is the situation that the body is attached to the rest of the boundary such that one has to make a compromise in the definition of the functions $\boldsymbol{w}_{\mathrm{drag}}$ and $\boldsymbol{w}_{\mathrm{lift}}$. For instance, in John (2002), the function $\boldsymbol{w}_{\mathrm{drag}}$ was defined as finite element function with the same polynomial degree as the finite element velocity. All degrees of freedom of $\boldsymbol{w}_{\mathrm{drag}}$ on $\overline{\Gamma_{\mathrm{body}}}$, i.e., also the degrees of freedom that are on the connections of the cylinder to the left and right boundary of the domain in Fig. D.6, were set to $\boldsymbol{w}_{\mathrm{drag}} = (1, 0, 0)^T$ and all other degrees of freedom to $\boldsymbol{w}_{\mathrm{drag}} = \boldsymbol{0}$. The same construction was used for $\boldsymbol{w}_{\mathrm{lift}}$

Table D.1 Example D.6

Coefficient	Schäfer and Turek (1996)	John (2002)
c_{drag}	$[6.05, 6.25]$	6.1853329
c_{lift}	$[0.008, 0.01]$	0.0094009839
Δp	$[0.165, 0.175]$	0.1708

Reference intervals and reference values

with $\boldsymbol{w}_{\text{lift}} = (0, 1, 0)^T$ for the degrees of freedom on $\overline{\Gamma_{\text{body}}}$. Then, formulas (D.5) and (D.8) were used with $A = d \cdot H = 0.041\,\text{m}^2$ and $V = U_{\text{mean}}$, leading to

$$c_{\text{drag}} = -\frac{500}{0.41}\big((\nu\nabla\boldsymbol{u}, \nabla\boldsymbol{w}_{\text{drag}}) + n(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{w}_{\text{drag}}) - (\nabla\cdot\boldsymbol{w}_{\text{drag}}, p)\big),$$
$$c_{\text{lift}} = -\frac{500}{0.41}\big((\nu\nabla\boldsymbol{u}, \nabla\boldsymbol{w}_{\text{lift}}) + n(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{w}_{\text{lift}}) - (\nabla\cdot\boldsymbol{w}_{\text{lift}}, p)\big).$$

Using quadrature rules for the evaluation of c_{drag} and c_{lift} with interior quadrature points, like Gaussian quadrature, the modification of $\boldsymbol{w}_{\text{drag}}$ and $\boldsymbol{w}_{\text{lift}}$ does not enter directly the numerical quadrature (values at the boundary of mesh cells are not seen by the quadrature points) and thus the impact of this modification is negligible. Table D.1 provides reference intervals from Schäfer and Turek (1996) and some reference values from John (2002).

□

D.2 Examples for Laminar Time-Dependent Flow Problems

Example D.7 (A Time-Dependent Example with Prescribed Solution in the Unit Square and with Homogeneous Dirichlet Boundary Conditions) Let $\Omega = (0, 1)^2$. The prescribed solution of this example is

$$\boldsymbol{u} = 2\pi\sin(t)\begin{pmatrix}\sin^2(\pi x)\sin(\pi y)\cos(\pi y)\\ -\sin(\pi x)\cos(\pi x)\sin^2(\pi y)\end{pmatrix}, \quad p = 20\sin(t)\left(x^2y - \frac{1}{6}\right), \tag{D.24}$$

see Fig. D.7 for a snapshot of the solution. By definition, it is $\boldsymbol{u}|_\Gamma = \boldsymbol{0}$. For a given value ν of the viscosity, the right-hand side $\boldsymbol{f}$ is computed such that (D.24) satisfies the Navier–Stokes equations (7.1).

This test problem might serve for supporting error estimates. To include the increasing as well as the decreasing temporal regime in (D.24), the final time should be chosen to be sufficiently large, e.g., $T = 5$ as in de Frutos et al. (2016c). □

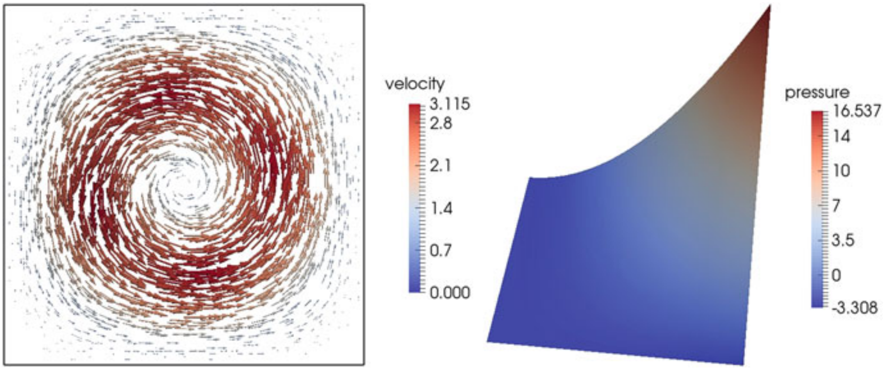

Fig. D.7 Example D.7. Computed solution at $t = 1.7$

Example D.8 (A Time-Dependent Flow Around a Cylinder at Re $= 100$ *with Steady-State Inflow)* This example describes a laminar Kármán vortex street. Its setup is very similar to Example D.5: the same domain, the same viscosity, the same boundary conditions at the cylinder, the lower and the upper wall, the do-nothing boundary condition (D.13) at the outlet, and also $\boldsymbol{f} = \mathbf{0}$. Only the inflow is five times as fast as in (D.12), given by

$$\boldsymbol{v}(0\,\mathrm{m}, y) = \frac{1}{0.41^2}\begin{pmatrix} 6y(0.41 - y) \\ 0 \end{pmatrix}\ {}^{\mathrm{m}}\!/_{\mathrm{s}}, \quad 0\,\mathrm{m} \le y \le 0.41\,\mathrm{m}. \tag{D.25}$$

Thus, the characteristic velocity scale and the Reynolds number multiply also with this factor, compared with Example D.5, and they become $U_{\mathrm{mean}} = 1\,{}^{\mathrm{m}}\!/_{\mathrm{s}}$ and $\mathrm{Re} = 100$, respectively. It turns out that the problem does not admit a stable steady-state solution in this situation, such that one has to consider as model the time-dependent incompressible Navier–Stokes equations (7.1). Using the same characteristic scales as in Example D.5 leads to the dimensionless viscosity $\nu = 10^{-3}$ in (7.1).

Describing an initial condition becomes necessary. There is no analytical expression for this condition. Instead, one starts a simulation with some initial condition and one lets the simulation run. Eventually, a Kármán vortex street develops, which finally turns out to be periodic. Having reached this state, the computed solution at some discrete time is saved and this solution will be used as initial condition for further simulations.

Analogously to Example D.5, the drag and lift coefficient at the cylinder and the pressure difference between the front and the back of the cylinder are quantities of interest. Similarly to the steady-state problem from Example D.5, it is possible to derive formulations with volume integrals. The only difference is that one has to take into account that the momentum equation possesses a term with the temporal derivative of the velocity. From (D.5) and (D.8), one obtains

$$c_{\text{drag}} = -20\big((\partial_t \boldsymbol{u}, \boldsymbol{w}_{\text{drag}}) + (\nu \nabla \boldsymbol{u}, \nabla \boldsymbol{w}_{\text{drag}}) + n(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{w}_{\text{drag}}) - (\nabla \cdot \boldsymbol{w}_{\text{drag}}, p)\big), \tag{D.26}$$

$$c_{\text{lift}} = -20\big((\partial_t \boldsymbol{u}, \boldsymbol{w}_{\text{lift}}) + (\nu \nabla \boldsymbol{u}, \nabla \boldsymbol{w}_{\text{lift}}) + n(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{w}_{\text{lift}}) - (\nabla \cdot \boldsymbol{w}_{\text{lift}}, p)\big), \tag{D.27}$$

where the definitions of $\boldsymbol{w}_{\text{drag}}$ and $\boldsymbol{w}_{\text{lift}}$ are given in Example D.5. A way for approximating the first terms consists in applying a backward difference, e.g.,

$$(\partial_t \boldsymbol{u}, \boldsymbol{w}_{\text{drag}}) \approx \frac{1}{\Delta t_{n+1}} \left((\boldsymbol{u}_{n+1} - \boldsymbol{u}_n) , \boldsymbol{w}_{\text{drag}} \right).$$

The pressure drop is computed according to (D.18). Typical evolutions of the drag coefficient, the lift coefficient, and the pressure drop are depicted in Fig. D.8.

Another functional of interest is the Strouhal number (2.23), which is closely connected to the length of the period (T^* in (2.23) is the length of the period). In

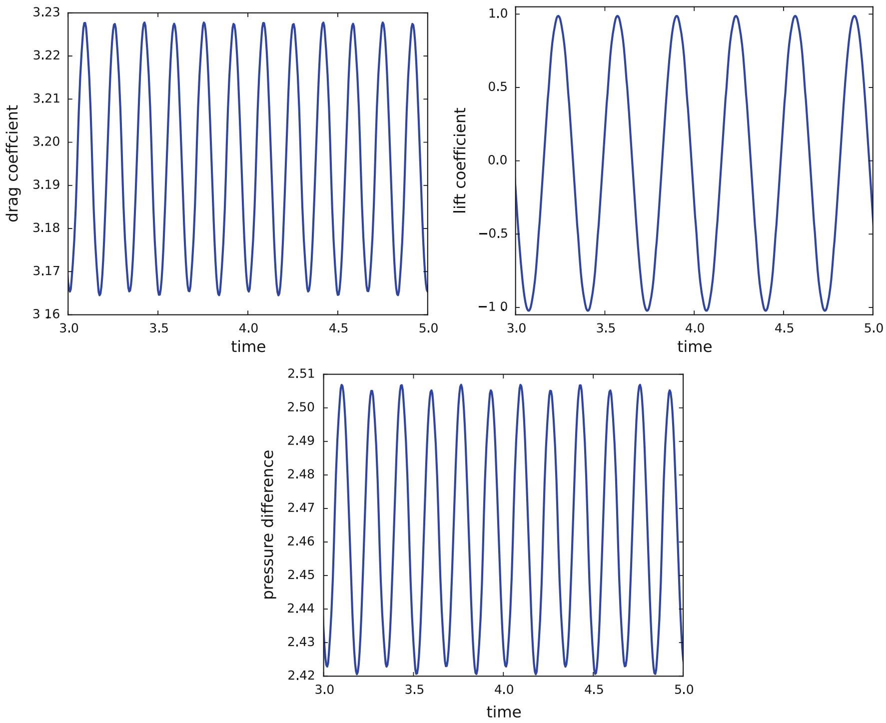

Fig. D.8 Example D.8. Temporal evolution of the quantities of interest

Table D.2 Example D.8

Coefficient	Reference interval
$c_{\text{drag,max}}$	$[3.22, 3.24]$
$c_{\text{lift,max}}$	$[0.98, 1.02]$
$\Delta p(t_{0,\text{per}} + T_{\text{per}}/2)$	$[2.46, 2.50]$
Strouhal number	$[0.295, 0.305]$

Reference intervals from Schäfer and Turek (1996). Here, $t_{0,\text{per}}$ is the start of a period (maximal value of the lift coefficient) and T_{per} is the length of a period

practice, one can compute the length of the period by considering the extrema of the drag or lift coefficient or the roots of the lift coefficient. Simulating the flow for a while, one takes the average length of a few subsequent periods.

Table D.2 presents reference intervals for the quantities of interest from Schäfer and Turek (1996). □

Example D.9 (A Time-Dependent Flow Around a Cylinder at Re $\in [0, 100]$ *with Time-Dependent Inflow)* This example describes the start and the decay of a vortex shedding. Its setup is almost the same as that of Example D.5. Only the inflow boundary condition is different. It is given by

$$\boldsymbol{v}(0\,\text{m}, y) = \sin\left(\frac{\phi t'}{8}\right) \frac{1}{0.41^2} \begin{pmatrix} 6y(0.41 - y) \\ 0 \end{pmatrix} {}^{\text{m}}\!/\!_{\text{s}}, \tag{D.28}$$

with $y \in [0, 0.41]\,\text{m}$, $t' \in [0, 8]\,\text{s}$, and the frequency $\phi = \pi\ {}^{1}\!/\!_{\text{s}}$. One obtains for the mean inflow in this example

$$U_{\text{mean}}(t') = \sin\left(\frac{\phi t'}{8}\right) \frac{1}{0.41^2} \frac{\int_0^{0.41} 6y(0.41 - y)\, dy}{\int_0^{0.41} dy} {}^{\text{m}}\!/\!_{\text{s}} = \sin\left(\frac{\phi t'}{8}\right) {}^{\text{m}}\!/\!_{\text{s}}.$$

Based on the diameter of $d = 0.1\,\text{m}$ of the cylinder, the mean inflow, and the kinematic viscosity μ/ρ, the Reynolds number of this flow $0 \le \text{Re}(t') \le 100$. For the boundary conditions at the outlet, one can use either Dirichlet boundary conditions that correspond to (D.28)

$$\boldsymbol{v}(2.2\,\text{m}, y) = \sin\left(\frac{\phi t'}{8}\right) \frac{1}{0.41^2} \begin{pmatrix} 6y(0.41 - y) \\ 0 \end{pmatrix} {}^{\text{m}}\!/\!_{\text{s}}, \tag{D.29}$$

$y \in [0, 0.41]\,\text{m}$, $t' \in [0, 8]\,\text{s}$, or do-nothing conditions

$$\mathbb{S}(t')\boldsymbol{n} = \left(2\mu\mathbb{D}(\boldsymbol{v})(t') - P(t')\mathbb{I}\right)\boldsymbol{n} = \boldsymbol{0}\,{}^{\text{N}}\!/\!_{\text{m}^2} \quad \text{on } \Gamma_{\text{out}},\ t' \in [0, 8]\,\text{s}. \tag{D.30}$$

The initial condition is set to be $\boldsymbol{v}_0 = \boldsymbol{0}\,{}^{\text{m}}\!/\!_{\text{s}}$. A visualization of the flow field at different times is presented in Fig. D.9.

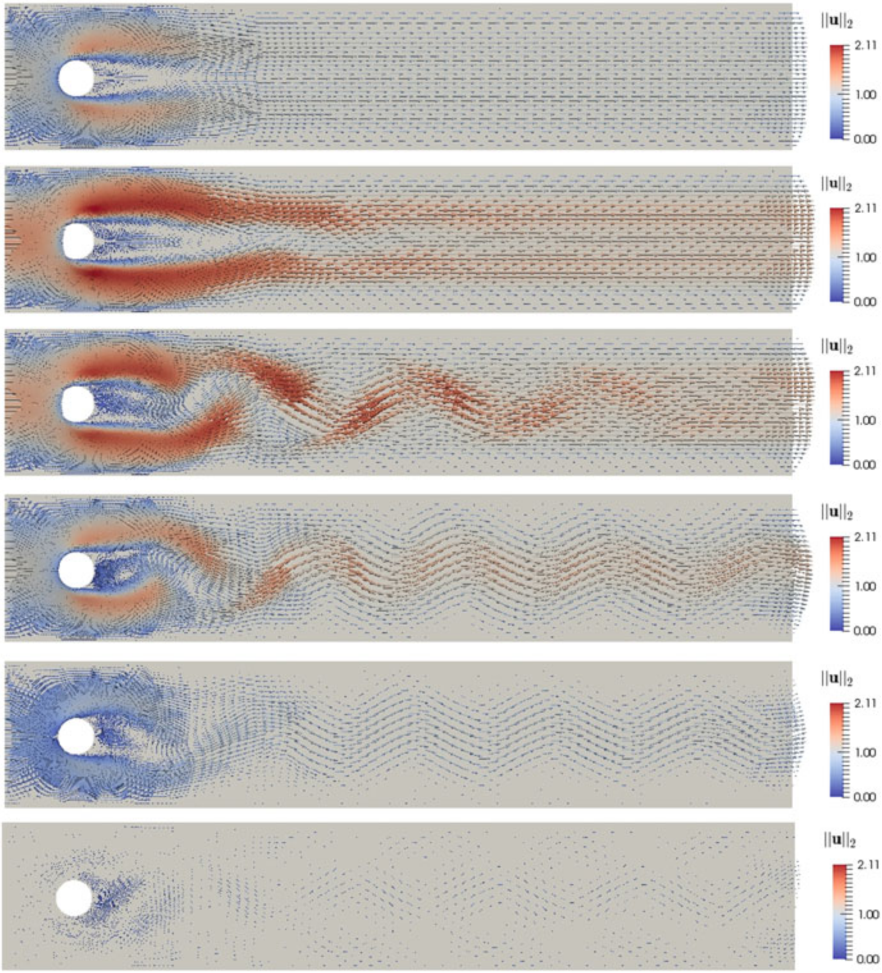

Fig. D.9 Example D.9. Temporal evolution of the velocity field, snapshots at 2, 4, 5, 6, 7, and 8 s

The inflow condition in dimensionless form is given by

$$\boldsymbol{u}(0, y) = \sin\left(\frac{\pi t}{8}\right) 0.41^{-2} \begin{pmatrix} 6y(0.41 - y) \\ 0 \end{pmatrix}, \quad 0 \leq y \leq 0.41,\ t \in [0, 8].$$

Like in Example D.5, quantities of interest are the drag coefficient (D.3), the lift coefficient (D.6), and the difference of the pressure between the front and the back of the cylinder

$$\Delta P(t) = P(t, (0.15, 0.2)) - P(t, (0.25, 0.2)).$$

Since the boundary of the circle cannot be represented exactly with a finite element triangulation, drag and lift should be computed with the area formulas (D.5) and (D.8) instead with the line formulas (D.4) or (D.7). In these formulas, one takes $U = 1\,{}^{\mathrm{m}}\!/\!_{\mathrm{s}}$, $L = 1$ m, $A = 0.1$ m, and $V = U_{\mathrm{mean}}(4\,\mathrm{s}) = 1\,{}^{\mathrm{m}}\!/\!_{\mathrm{s}}$. Then, one obtains the formulas (D.26) and (D.27).

For the quantities of interest, accurate reference curves were computed in John and Rang (2010), see Fig. D.10. These curves were obtained for the case that the Dirichlet boundary condition (D.29) at the outlet was applied. Reference values are provided in Table D.3. □

Example D.10 (Exponentially Decaying Flows in Two and Three Dimensions, Beltrami Flows) Let $\Omega \subset \mathbb{R}^3$ be a domain and consider a family of velocity and

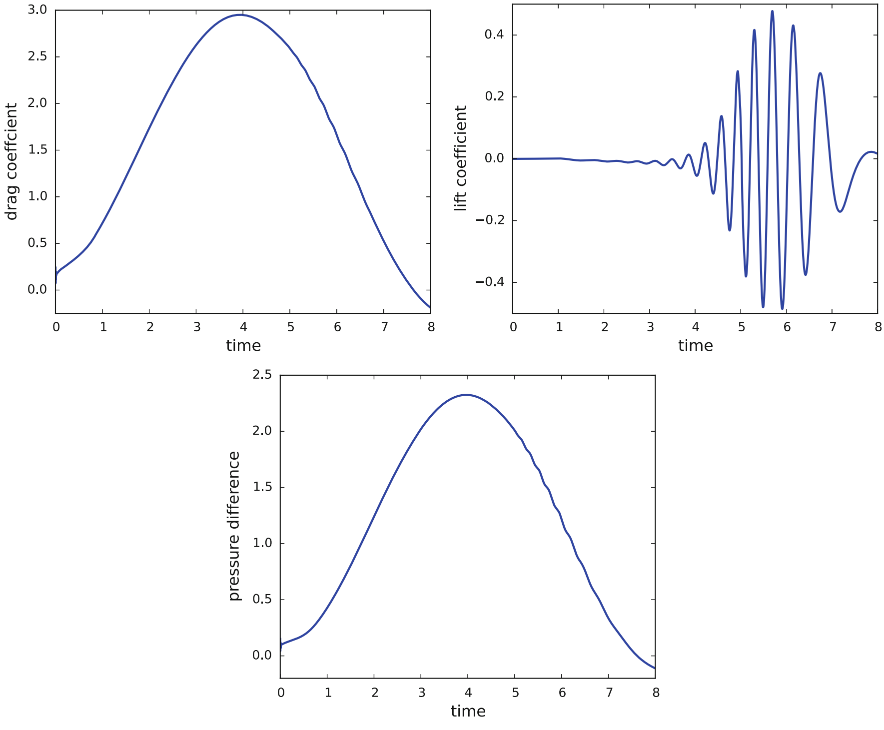

Fig. D.10 Example D.9. Reference curves for the quantities of interest from John and Rang (2010)

Table D.3 Example D.9

Coefficient	Value	Time
$c_{\mathrm{drag,max}}^{\mathrm{ref}}$	2.950918381	3.93625
$c_{\mathrm{lift,max}}^{\mathrm{ref}}$	0.47787543	5.69250
$\Delta p(8)^{\mathrm{ref}}$	−0.11161567	

Reference values from John and Rang (2010) pressure fields

$$\boldsymbol{u} = -\alpha \exp\left(-\nu\gamma^2 t\right) \begin{pmatrix} \exp(\alpha x)\sin(\alpha y \pm \gamma z) + \exp(\alpha z)\cos(\alpha x \pm \gamma y) \\ \exp(\alpha y)\sin(\alpha z \pm \gamma x) + \exp(\alpha x)\cos(\alpha y \pm \gamma z) \\ \exp(\alpha z)\sin(\alpha x \pm \gamma y) + \exp(\alpha y)\cos(\alpha z \pm \gamma x) \end{pmatrix},$$

$$\begin{aligned} p = &-\frac{\alpha^2}{2}\exp\left(-2\nu\gamma^2 t\right)\Big[\exp(2\alpha x) + \exp(2\alpha y) + \exp(2\alpha z) \\ &+2\sin(\alpha x \pm \gamma y)\cos(\alpha z \pm \gamma x)\exp(\alpha(y+z)) \\ &+2\sin(\alpha y \pm \gamma z)\cos(\alpha x \pm \gamma y)\exp(\alpha(z+x)) \\ &+2\sin(\alpha z \pm \gamma x)\cos(\alpha y \pm \gamma z)\exp(\alpha(x+y))\Big] + C, \end{aligned} \tag{D.31}$$

where $\alpha, \gamma \in \mathbb{R}$ are user-chosen parameters, see Fig. D.11, and C has to be chosen, in case of imposing Dirichlet boundary conditions at $\partial\Omega$, such that the integral mean of p vanishes. Inserting (D.31) in the Navier–Stokes equations (7.1) reveals with a direct calculation that $\boldsymbol{f} = \boldsymbol{0}$. The derivation of this family of solutions was performed in Ethier and Steinman (1994). It was based on the following principles:

- the velocity should be divergence-free,
- the temporal derivative should balance the viscous term $\partial_t \boldsymbol{u} - \nu\Delta\boldsymbol{u} = \boldsymbol{0}$,
- the convective term can be expressed by the gradient of a scalar function, which is the negative of the gradient of the pressure, $(\boldsymbol{u}\cdot\nabla)\boldsymbol{u} + \nabla p = \boldsymbol{0}$.

A straightforward calculation shows that

$$\nabla \times (\nabla \times \boldsymbol{u}) = \boldsymbol{0}. \tag{D.32}$$

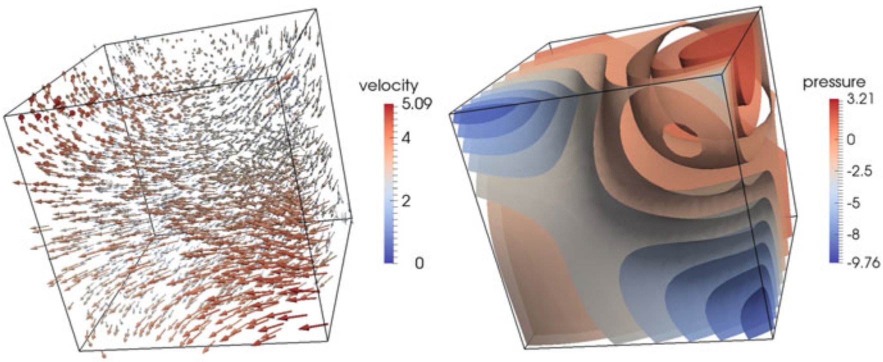

Fig. D.11 Example D.10. $\Omega = (-1,1) \times (-1,1) \times (0,2)$, velocity and pressure at $t = 0$ for $\alpha = \pi/4$ and $\gamma = \pi/2$

Flows with property (D.32) are called Beltrami flows.

A two-dimensional exponentially decaying flow problem in $\Omega = (0, \pi)^2$ was already proposed in Taylor (1923)

$$\boldsymbol{u} = \exp\left(-2\nu\gamma^2 t\right)\begin{pmatrix} -\cos(\gamma x)\sin(\gamma y) \\ \sin(\gamma x)\cos(\gamma y)\end{pmatrix},$$

$$p = -\frac{1}{4}\exp\left(-4\nu\gamma^2 t\right)\left[\cos(2\gamma x) + \cos(2\gamma y)\right].$$

Numerical studies with a solution of this type can be found, e.g., in Chorin (1968).
□

Remark D.11 (A Time-Dependent Flow Around a Cylinder in Three Dimensions at Re $\in [0, 100]$ *with Time-Dependent Inflow)* A three-dimensional extension of Example D.9 was also defined in Schäfer and Turek (1996) whose definition is similar to Example D.6. In comparison with the latter example, the inflow is time-dependent and in the most part of the time interval it is stronger. However, in John (2006), it was observed that in the proposed example there is no vortex shedding. The absence of this feature reduces the interest in using this example. □

D.3 Examples for Turbulent Flow Problems

Example D.12 (Turbulent Channel Flows) Turbulent channel flow problems are governed by the dimensionless incompressible Navier–Stokes equations of the form

$$\begin{aligned} \partial_t \boldsymbol{u} - 2\nabla\cdot\left(\mathrm{Re}_\tau^{-1}\mathbb{D}(\boldsymbol{u})\right) + (\boldsymbol{u}\cdot\nabla)\,\boldsymbol{u} + \nabla p &= \boldsymbol{f} \text{ in } (0,T]\times\Omega, \\ \nabla\cdot\boldsymbol{u} &= 0 \text{ in } (0,T]\times\Omega, \end{aligned} \tag{D.33}$$

with $\boldsymbol{f} = (1, 0, 0)^T$. The dimensionless equations (D.33) were derived with the characteristic scales $L = 1\,\mathrm{m}$ and $U = 1\,\mathrm{m/s}$. It follows that the dimensionless viscosity is $\mathrm{Re}^{-1} = \nu = \mathrm{Re}_\tau^{-1}$, see Remark 2.19. In addition, one obtains with (2.23) that the kinematic viscosity of the fluid is $\nu = \mathrm{Re}_\tau^{-1}\,\mathrm{m^2/s}$.

There are benchmark problems for several values of Re_τ.

The Turbulent Channel Flow at $Re_\tau = 180$ The domain of this problem is given by

$$\Omega = (-2\pi,\, 2\pi)\times(0,\, 2H)\times\left(-\frac{2}{3}\pi,\, \frac{2}{3}\pi\right),$$

with $H = 1$ being the channel half width and $\mathrm{Re}_\tau = 180$ being the Reynolds number based on the channel half width, the kinematic viscosity ν of the fluid, and the shear or friction velocity U_τ, see Remark 8.13. It follows for the friction velocity that

$$\mathrm{Re}_\tau = \frac{U_\tau H}{\nu} \implies U_\tau = \frac{\mathrm{Re}_\tau \nu}{H} = \frac{\mathrm{Re}_\tau \mathrm{Re}_\tau^{-1}}{H} = 1\,{}^{\mathrm{m}}\!/\!{}_{\mathrm{s}} \implies u_\tau = \frac{U_\tau}{U} = 1.$$

There are periodic conditions in the streamwise x- and the spanwise z-direction and no-slip conditions $\boldsymbol{u} = \mathbf{0}$ at the solid walls at $y = 0$ and $y = 2$.

The definition of an initial condition might be based on the known mean velocity profile $U_{\mathrm{mean}}(y)$ from the data file `chan180.means` provided in Moser et al. (1999), see Fig. D.12. This mean velocity profile is interpolated linearly and noise is added, e.g., in the form as proposed in Gravemeier (2006b),

$$\begin{aligned}
\boldsymbol{u}_1(0; x, y, z) &= U_{\mathrm{mean}}(y) + 0.1\, U_{\mathrm{bulk}}\, \psi(x, y, z),\\
\boldsymbol{u}_2(0; x, y, z) &= 0.1\, U_{\mathrm{bulk}}\, \psi(x, y, z),\\
\boldsymbol{u}_3(0; x, y, z) &= 0.1\, U_{\mathrm{bulk}}\, \psi(x, y, z).
\end{aligned} \tag{D.34}$$

Here,

$$U_{\mathrm{bulk}} = \frac{1}{H} \int_0^H U_{\mathrm{mean}}(y)\, dy \approx 15.6803, \tag{D.35}$$

where $U_{\mathrm{mean}}(y)$ was interpolated by a cubic spline for obtaining this value by numerical integration. As noise often a random function of the form (using `C++`–notation)

$$\psi(x, y, z) = \frac{2\,\mathrm{rand}()}{\mathrm{RAND\_MAX}()} - 1 \;\in\; [-1, 1]$$

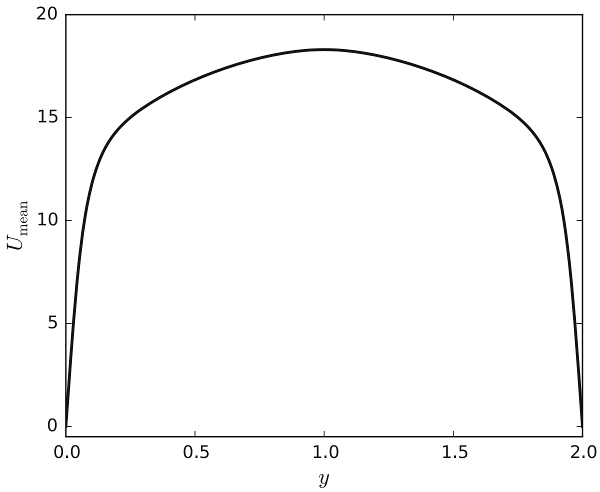

Fig. D.12 Example D.12. Mean velocity profile $U_{\mathrm{mean}}(y)$ for the turbulent channel flow at $\mathrm{Re}_\tau = 180$ from the data file `chan180.means` provided in Moser et al. (1999)

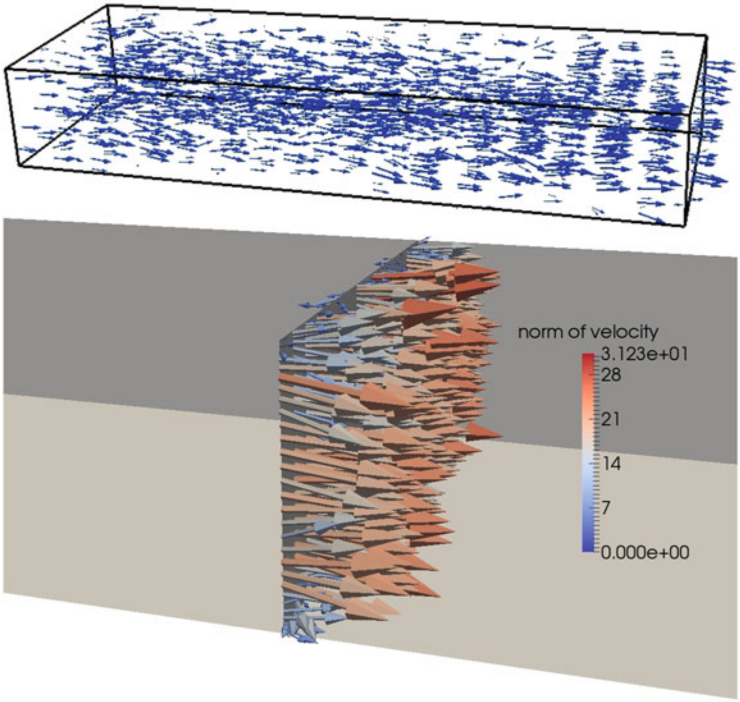

Fig. D.13 Example D.12. Snapshot of the velocity computed in a turbulent channel flow problem with $\mathrm{Re}_\tau = 180$, whole channel (*top*) and at a cut plane perpendicular to the x-direction (*bottom*, plane $y = 0$ in *light gray*, $z = 2\pi/3$ in *dark gray*)

is used. Altogether, the initial velocity field (D.34) is obtained by disturbing the interpolated mean velocity profile $U_{\mathrm{mean}}(y)$ by a random velocity fluctuation of 20 % of the bulk velocity U_{bulk} (10 % in negative and 10 % in positive direction). A snapshot of the computed flow field is presented in Fig. D.13. It can be seen that the flow is in the mean from left to right. The velocity at the cut plan reveals that the flow is in the mean faster in the interior of the channel than close to the boundary, which is expected. However, close to the boundary, there are eddies such that also a flow in the opposite direction occurs.

The Turbulent Channel Flow at $Re_\tau = 395$ This problem is given in

$$\Omega = (-\pi, \pi) \times (0, 2) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right),$$

again with the solid walls at $y = 0$ and $y = 2$ and periodic boundary conditions at the other boundaries. In the same way as in the case $\mathrm{Re}_\tau = 180$, one finds that $u_\tau = U_\tau / U = 1$. With the reference mean velocity from Moser et al. (1999), one obtains for the bulk velocity $U_{\mathrm{bulk}} = 17.5452$. The initial condition can be defined the same way as for $\mathrm{Re}_\tau = 180$.

Statistics of Interest Let $\langle\cdot\rangle_t$ and $\langle\cdot\rangle_s$ denote averages in time and in the spatial direction(s) of homogeneity. These directions are for turbulent channel flow problems the streamwise x-direction and the spanwise z-direction. Using an equidistant time step, the average in time, e.g., of the finite element velocity can be computed with the arithmetic mean

$$\langle \boldsymbol{u}^h(t,\boldsymbol{x})\rangle_t = \frac{1}{N_t+1}\sum_{n=0}^{N_t} \boldsymbol{u}^h(t_n,\boldsymbol{x}),$$

where N_t is the number of time steps. Likewise, if equidistant grids are used in the directions of homogeneity, the arithmetic mean can be applied for computing the spatial average, e.g.,

$$\langle \boldsymbol{u}^h(t_n,x,y,z)\rangle_s = \frac{1}{N_x}\frac{1}{N_z}\sum_{i=1}^{N_x}\sum_{j=1}^{N_z} \boldsymbol{u}^h(t_n,x_i,y,z_j),$$

where N_x and N_z are the numbers of degrees of freedom in the x- and z-direction, respectively.

Statistics of interest include the mean velocity and the Reynolds stresses, which are defined as the first order averaged quantity

$$\boldsymbol{u}^h_{\text{mean}}(y) = \langle\langle \boldsymbol{u}^h(t,x,y,z)\rangle_s\rangle_t$$

and the second order averaged quantities

$$\mathbb{T}^h_{ij,\text{mean}} = \langle\langle u^h_i u^h_j\rangle_s\rangle_t - \langle\langle u^h_i\rangle_s\rangle_t\langle\langle u^h_j\rangle_s\rangle_t, \quad i,j=1,2,3, \tag{D.36}$$

see Remark 8.32 for the definition of the Reynolds stress tensor. Concerning the mean velocity, in particular the first component $u^h_{1,\text{mean}}(y)$ of $\boldsymbol{u}^h_{\text{mean}}(y)$ is of interest.

The definition of the Reynolds stress is not unique in the literature, see the discussion in John and Roland (2007). It was pointed out in Winckelmans et al. (2002) that the diagonal Reynolds stresses, i.e., $i = j$ in (D.36), computed with the solution of turbulent flow simulations cannot be compared with the diagonal stresses of the turbulent flow field. The reason is that in the turbulence modeling the trace of the subgrid scale stress tensor is used to define a modified pressure, see Remark 8.65, and this trace is not available in the simulations. Only, the so-called rms (root mean squared) turbulence intensities

$$u^h_{i,\text{rms}} = \left|\mathbb{T}^h_{ii,\text{mean}} - \frac{1}{3}\sum_{j=1}^{3}\mathbb{T}^h_{jj,\text{mean}}\right|^{1/2} \qquad i=1,2,3, \tag{D.37}$$

can be studied directly.

In addition, it is pointed out in Winckelmans et al. (2002) that for the second order statistics (D.36) and (D.37), a contribution from the used turbulence model should be included. For some turbulence models, this contribution is rather easy to identify and to implement, e.g., for the Smagorinsky model from (8.62) and (8.66), see also John and Roland (2007). However, there are turbulence models where the inclusion of their contribution is difficult.

Aspects of the Discretization In simulations, usually uniform grids are used in x- and z-direction and graded grids in the wall-normal direction (y-direction), which become finer towards the walls, e.g., see Gravemeier (2006a,b) for two proposals. The length of the (equidistant) time step has to be chosen sufficiently small to resolve the temporal dynamics of the flow. Studies in Choi and Moin (1994) suggest to choose

$$\Delta t^{+} = u_\tau \mathrm{Re}_\tau \Delta t \le 0.4. \tag{D.38}$$

An Aspect of Finite Element Simulations Since the flow is incompressible, the bulk velocity should be constant during the simulation. However, finite element functions are in general not weakly divergence-free, but only discretely divergence-free. Thus, a finite element discretization cannot be expected to lead automatically to a conservation of the bulk velocity, which corresponds to the conservation of mass. A possibility to account for the difference of the computed bulk velocity and (D.35) consists in a dynamical adjustment of the right-hand side of the Navier–Stokes equations as proposed in John and Roland (2007). The flow is driven by a pressure gradient. Let $U_{\mathrm{bulk,sim}}(t_n)$ be the bulk velocity of the computed solution at time t_n. Then, the right-hand side of the Navier–Stokes equations (D.33) at t_{n+1} is defined by

$$\boldsymbol{f} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \frac{1}{\Delta t_n} \begin{pmatrix} U_{\mathrm{bulk}} - U_{\mathrm{bulk,sim}}(t_n) \\ 0 \\ 0 \end{pmatrix}, \tag{D.39}$$

where Δt_n is the length of the time step. Note that for quantities with physical dimensions, the correction term in (D.39) possesses the correct physical unit. Thus, if $U_{\mathrm{bulk,sim}}(t_n) < U_{\mathrm{bulk}}$, the flow will be accelerated which leads to an increase of the bulk velocity of the computed solution. In the case $U_{\mathrm{bulk,sim}}(t_n) > U_{\mathrm{bulk}}$, the forcing becomes smaller than 1 and $U_{\mathrm{bulk,sim}}$ becomes smaller in the next discrete time.

It is reported in John and Roland (2007) that if the dynamical adjustment of the driving force (D.39) is not applied, as well an increase as a decrease of the bulk velocity of the computed solution could be observed, depending on the turbulence model. Using (D.39), the bulk velocity still showed some oscillations but they stayed always close (differences in general less than 1 %) to the value given in (D.35). □

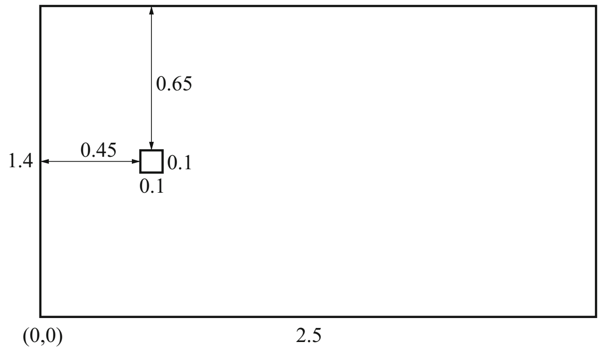

Fig. D.14 Example D.13. Cross section (x-y plane) of the domain (all length in m), the height of the channel is $H = 0.4\,\mathrm{m}$

Example D.13 (A Turbulent Flow Around a Cylinder with Square-Shaped Cross Section at Re $=$ 22000*)* This example, describing a channel flow around an obstacle, was defined in Rodi et al. (1997). Even if the obstacle is actually a column, it is usually called a cylinder in the literature. Figure D.14 presents a sketch of the cross section of the flow domain, where the height of the channel is $H = 0.4\,\mathrm{m}$. The inflow is prescribed by

$$\boldsymbol{u}(t, 0, y, z) = \begin{pmatrix} 1 + 0.04\,\mathtt{rand} \\ 0 \\ 0 \end{pmatrix}, \tag{D.40}$$

where rand is a random function with values in $[-0.5, 0.5]$. The noise in the inflow serves to stimulate the turbulence. No-slip boundary conditions are prescribed at the column. Outflow boundary conditions (2.37) are set at $x = 2.5\,\mathrm{m}$. On all other boundaries, free slip conditions (2.34) are used. Since the initial condition is not known, one starts the simulations with some condition, e.g., the interpolation of the inlet condition at the initial time into the domain, and performs the simulations until the flow has been developed, before the statistics of interest are computed. There are no external forces acting on the flow. This example describes a statistically periodic flow, see Fig. D.15 for a snapshot.

The Reynolds number of the flow, based on the mean inflow $U = 1\,\mathrm{m/s}$, the side length of the cylinder $D = 0.1\,\mathrm{m}$, and the viscosity $\nu = 1/220000\,\mathrm{m^2/s}$ is $\mathrm{Re} = UD/\nu = 22000$.

As usual for flows around bodies, one is interested in the drag and the lift coefficient. These coefficients can be computed as volume integrals, see (D.5) and (D.8) for the formulas with dimensionless quantities. For this example, the

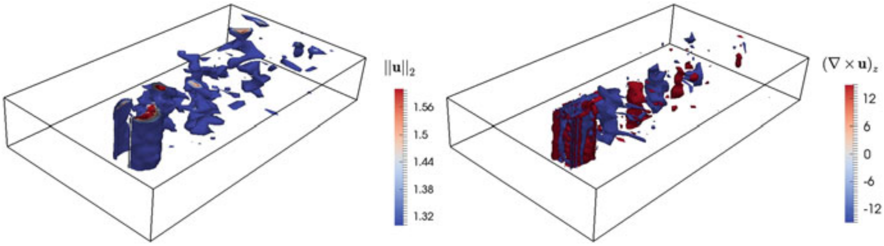

Fig. D.15 Example D.13. Snapshot of isosurfaces of the velocity magnitude $\|\boldsymbol{u}\|_2$ (*left*) and the third component of $\nabla \times \boldsymbol{u}$, see (3.151), (*right*). The vortex shedding behind the cylinder is clearly visible

parameters in these formulas are

$$L = 1\,\text{m}, \quad V = U, \quad A = DH, \quad \boldsymbol{w}_{\text{drag}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{w}_{\text{lift}} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

At the connections of the cylinder and the free slip boundaries, the vectors $\boldsymbol{w}_{\text{drag}}$ and $\boldsymbol{w}_{\text{lift}}$ should be set as described in Example D.6. The Strouhal number is defined by

$$\text{St} = \frac{DU}{T},$$

where T is the average length of a period. Quantities of interest are the Strouhal number, time-averaged drag and lift coefficients, $\overline{c}_{\text{drag}}$ and $\overline{c}_{\text{lift}}$, and root mean squared (rms) values for c_{drag}, c_{lift}, which are defined by

$$c_{\text{drag,rms}} = \left(\frac{1}{N_t} \sum_{n=1}^{N_t} \left(c_{\text{drag}}(t_n) - \overline{c}_{\text{drag}} \right)^2 \right)^{1/2},$$

$$c_{\text{lift,rms}} = \left(\frac{1}{N_t} \sum_{n=1}^{N_t} \left(c_{\text{lift}}(t_n) - \overline{c}_{\text{lift}} \right)^2 \right)^{1/2},$$

where the summation covers all discrete times in the time interval for which $\overline{c}_{\text{drag}}, \overline{c}_{\text{lift}}$ are computed. Some experimental values are provided in Rodi et al. (1997)

$$\overline{c}_{\text{drag}} \in [1.9, 2.1], \quad c_{\text{drag,rms}} \in [0.1, 0.2], \quad c_{\text{lift,rms}} \in [0.7, 1.4], \quad \text{St} = 0.132.$$

Because of the statistically periodic vortex shedding, it is $\overline{c}_{\text{lift}} = 0$. □

Example D.14 (A Turbulent Flow Around a Wall-Mounted Cube at $Re = 40000$*)* In this example, the fluid flows not only around a body but also across that body. It is defined, e.g., in Hoffman (2005), Hoffman and Johnson (2006). An example of this type was also proposed in Rodi et al. (1997). The bottom of the domain, at $z = 0$, is sketched in Fig. D.16. The height of the cube is H and the height of the channel $2H$. In Hoffman (2005), Hoffman and Johnson (2006), simulations for $H = 0.1$ m were presented.

There are no body forces in this example, i.e., $\boldsymbol{f} = \boldsymbol{0}$. In Table D.4, nodal values of the mean velocity in streamwise direction are provided. To obtain an inlet condition for simulations, these values can be interpolated to the degrees of freedom on the inlet boundary and some noise can be added. In this way, one obtains an inlet condition of form (D.40) with 1 replaced by the interpolation of $u_{1,\mathrm{mean}}$ and possibly

Fig. D.16 Example D.14. *Bottom* plane $z = 0$ of the domain

Table D.4 Example D.14

Point	$u_{1,\mathrm{mean}}$	Point	$u_{1,\mathrm{mean}}$	Point	$u_{1,\mathrm{mean}}$	Point	$u_{1,\mathrm{mean}}$
0	0						
1	0.2166	17	1.0315	33	1.0930	49	1.0083
2	0.6292	18	1.0461	34	1.0928	50	0.9871
3	0.6926	19	1.0587	35	1.0924	51	0.9654
4	0.7307	20	1.0661	36	1.0921	52	0.9445
5	0.7671	21	1.0723	37	1.0911	53	0.9229
6	0.7941	22	1.0770	38	1.0898	54	0.9029
7	0.8191	23	1.0813	39	1.0885	55	0.8814
8	0.8466	24	1.0843	40	1.0869	56	0.8622
9	0.8716	25	1.0868	41	1.0835	57	0.8360
10	0.8931	26	1.0890	42	1.0796	58	0.8091
11	0.9168	27	1.0890	43	1.0747	59	0.7851
12	0.9367	28	1.0890	44	1.0679	60	0.7582
13	0.9586	29	1.0890	45	1.0606	61	0.7199
14	0.9787	30	1.0900	46	1.0527	62	0.6857
15	0.9986	31	1.0930	47	1.0388	63	0.4141
16	1.0153	32	1.0930	48	1.0222	64	0

Mean values for the inlet condition at 65 equally distributed points (from J. Hoffman, personal communication), point 0 corresponds to $z = 0$, point 64 corresponds to $z = 2H$

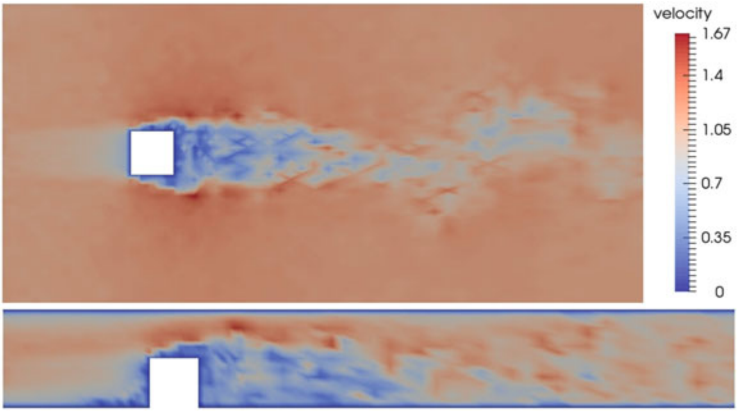

Fig. D.17 Example D.14. Snapshot of the velocity magnitude $\|\boldsymbol{u}\|_2$ in the cut planes $z = H$ (*top*) and $y = 3.5H$ (*bottom*)

with some other factor in front of the random term (the simulations for Fig. D.17 were performed with the factor 0.1). There are no-slip boundary conditions at the cube, the lower, and the upper boundary. On the lateral boundaries, free slip conditions (2.34) are prescribed. At the outflow, the do-nothing condition (2.37) can be used.

An initial condition is not known. One can start the simulations by interpolating the inflow condition at the initial time into the domain, setting the no-slip boundary conditions at all boundaries that are equipped with these conditions, and letting the simulations run until the flow field is developed. In Hoffman (2005), Hoffman and Johnson (2006), it can be seen that this state is reached after 3–5 s.

The mean inlet velocity is around $1\,\mathrm{m/s}$. Thus, taking as characteristic scales for defining the Reynolds number $U = 1\,\mathrm{m/s}$, $H = 0.1\,\mathrm{m}$ and considering a fluid with kinematic viscosity $\nu = 2.5 \cdot 10^{-6}\,\mathrm{m^2/s}$ leads to $\mathrm{Re} = 40000$. Hence, the flow is turbulent. Deriving the dimensionless Navier–Stokes equations (7.1) with the characteristic scales $U = 1\,\mathrm{m/s}$ and $L = 1\,\mathrm{m}$ gives the dimensionless viscosity $\nu = 1/400000$ in (7.1). A snapshot of a simulation of the flow field is depicted in Fig. D.17.

In Hoffman (2005), Hoffman and Johnson (2006), the mean drag coefficient at the cube was in the focus of the numerical studies. Adaptively refined grids were used and the SUPG stabilization was applied, see the last term of (5.34) with $\partial_t \boldsymbol{u}$ to be added in the residual, which gives a so-called implicit LES model. The adaptive refinement was controlled with a DWR approach, compare Sect. 6.4. Having applied the time averaging in an interval of 4 s, the time-averaged drag coefficient $\overline{c}_{\mathrm{drag}} \approx 1.5$ was obtained. The flow turns out to be statistically periodic such that $\overline{c}_{\mathrm{lift}} = 0$. For evaluating the drag and lift coefficient, formulas (D.5) and (D.8) can be used with

$$L = 1\,\mathrm{m}, \quad V = U, \quad A = H^2, \quad \boldsymbol{w}_{\mathrm{drag}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{w}_{\mathrm{lift}} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

The boundary of the body is attached to the boundary of the channel. Thus, in the actual definitions of $\boldsymbol{w}_{\mathrm{drag}}$ and $\boldsymbol{w}_{\mathrm{lift}}$, one has to use a modification compared with Remark D.2, e.g., the modification that is described in Example D.6. □

Appendix E
Notations

	Symbol	Meaning	Sect.
A	$a(\cdot,\cdot)$	Bilinear form in saddle point problems	3.1
	$\boldsymbol{a}_i$	Vertices of a mesh cell	B.2
	$\boldsymbol{a}_{ij}$	Point of a mesh cell	B.2
	A	Operators in saddle Point problems	3.1
	A	Matrix for velocity-velocity coupling in the finite element system	4.3
	$\boldsymbol{A}$	Nonlinear viscous operator for Smagorinsky model	8.3.2
	$\mathscr{A}$	Global saddle point matrix	9
	$\mathscr{A}_j$	Local saddle point matrix	9.2.2
	$\mathcal{A}$	Attractor	8.1
	A^h	Discrete Stokes operator	7.2
	$\mathcal{A}_{\text{Ler}}$	Attractor for Leray-α model	8.6
	A_{pspg}	Bilinear form for PSPG method	4.5.1
	A_{spg}	Bilinear form for SUPG/PSPG/grad-div method	5.3.2
	A_U	Operator in two-scale VMS methods	8.8.1
	A_{wave}	Operator in turbulence models based on an approximation in wave number space	8.4
B	$\boldsymbol{b}$	Convection in Oseen problem	5.1
	$\boldsymbol{b}$	Vector in affine mappings	B.2
	$b(\cdot,\cdot)$	Bilinear form in abstract saddle point problem	3.1
	$b(\cdot,\cdot)$	Bilinear form coupling velocity and pressure space	3.2
	$b^h(\cdot,\cdot)$	Bilinear form coupling discrete velocity and pressure space	3.3
	B	Operators in saddle point problems	3.1
	B	Matrix for pressure-velocity coupling in the finite element system	4.3
	B^T	Matrix for velocity-pressure coupling in the finite element system	4.3
	B^h	Discrete divergence operator	3.3
	B_K	Matrix in affine mappings	B.2

V. John, *Finite Element Methods for Incompressible Flow Problems*, Springer Series in Computational Mathematics 51, DOI 10.1007/978-3-319-45750-5

C	c	Coefficient function in the Oseen problem	5.1
	c_0	Lower bound for coefficient function in the Oseen problem	5.1
	c_{drag}	Drag coefficient	D
	c_{lift}	Lift coefficient	D
	curl	Rotation or curl of a 2d function	3.7
	curl	Rotation or curl of a scalar function in 2d	3.7
	C_{inv}	Constant in inverse estimate	C.4
	C_{PF}	Constant in the Poincaré–Friedrichs inequality	A.2
	C	Matrix of pressure-pressure coupling	4.5.1
	C_{os}	Constant in error estimate for Oseen problem	5.2
	C_S	Smagorinsky coefficient	8.3.1
	$C^s(\Omega)$, $C^s(K)$, ...	Space of s times continuous differentiable functions on $\Omega, K, \ldots$	A.2
D	d	Dimension	
	$d(\cdot,\cdot)$	Metric, distance	A.1
	$d_{\text{H}}(X)$	Hausdorff dimension of a set X	8.1
	div	Divergence operator	3.2
	div^h	Discrete divergence operator	3.3
	$\mathbb{D}(\boldsymbol{v})$	Velocity deformation tensor	2.2
E	e	Error for primal problem	6.4
	e^*	Error for dual linearized problem	6.4
	$\boldsymbol{e}_i$	ith Cartesian coordinate vector	
	E	Face, edge	B.2
	$\mathcal{E}(K)$	Set of all $(d-1)$ dimensional faces of K	B.2
	$\mathcal{E}^h$	Set of all $(d-1)$ dimensional faces of a triangulation that are not at the boundary	3.3
	$\overline{\mathcal{E}}^h$	Set of all $(d-1)$ dimensional faces of a triangulation	3.3
	E_0	Imbedding operator	3.1
	$E(k)$	Kinetic energy spectrum	8.1
	$E(t,\boldsymbol{k})$	Kinetic energy for wave number $\boldsymbol{k}$ at time t	8.1
	$E(t,k)$	Kinetic energy for wave numbers $\\|\boldsymbol{k}\\|_2 = k$ at time t	8.1
F	$\boldsymbol{f}$	Dimensionless body force	2.2
	$\underline{f}$	Right-hand side vector of algebraic momentum equation	4.3
	$\underline{f_p}$	Right-hand side vector of algebraic continuity equation	4.5.1
	$\boldsymbol{f}_{\text{ext}}$	External force	2.2
	$\boldsymbol{f}_{\text{net}}$	Net force density	2.2
	$\boldsymbol{f}_{\boldsymbol{w}}$	Force exerted by the flow on a body in the direction $\boldsymbol{w}$	D
	F_K	Reference map to mesh cell K	B.2
	F_M	Reference map to macroelements	3.5.4
	$\mathcal{F}_{\hat{M}}$	Family of macroelements	3.5.4
	F_{drag}	Drag force	D
	F_{lift}	Lift force	D

G	$\boldsymbol{g}$	Dirichlet boundary condition	2.4
	g_δ	Gaussian filter	8.2.1
	g_{box}	Box filter	8.2.1
	grad	Gradient operator	3.2
	grad^h	Discrete gradient operator	3.3
	G_{conv}	Convolution operator	8.5
	G_{deconv}	Deconvolution operator	8.5
	$G_{\mathrm{deconv},N}$	Approximate deconvolution operator	8.5
H	h	Characteristic parameter of triangulation	C.1
	h_K	Diameter of a mesh cell K	C.1
	$h_{K,\mathrm{short}}$	Shortest edge of a mesh cell	8.3
	$h_{K,\mathrm{vol}}$	Cubic root of the volume of a mesh cell	8.3
	h_M	Diameter of a macro cell M	3.5
	H	Channel width or half width	D.2
I	I	Identity operator	
	$\mathbb{I}$	Identity tensor	2.2
	I^h	Interpolation operator in a finite element space	3.3
	I^h_{div}	Discrete divergence preserving interpolation operator in V^h	3.3
	I^h_{St}	Stokes projection	4.2
	$\mathcal{I}_l$	Set of degrees of freedom	9.2.2
J	$J(\cdot)$	Functional of interest	6.4
	J_0, J_1	Quadratic functionals	3.1
K	$\boldsymbol{k}$	Wave number vector	8.1
	k_c	Cut-off wave number	8.2
	K	Mesh cell	B.2
	$\hat{K}$	Reference mesh cell	B.2
	$\mathring{K}$	Interior of a mesh cell K	B.2
	$K_k(\underline{r}, \mathscr{A})$	Krylov subspace of dimension k spanned by $\underline{r}$ and $\mathscr{A}$	9.1
L	L	Highest geometric level	9.2.2
	$L(\cdot)$	Functional	6.4
	L_{pspg}	Linear form for PSPG method	4.5.1
	L_{spg}	Linear form for SUPG/PSPG/grad-div method	5.3.2
	$\mathcal{L}(V,Q)$	Space of linear operators $V \to Q$	3.1
M	$\boldsymbol{m}_E$	Barycenter of a face E	B.3
	M	Macroelement	3.5.4
	$\hat{M}$	Reference macroelement	3.5.4
	M_K	Macroelement containing the mesh cell K	3.3

N	$\boldsymbol{n}$	Outward pointing unit normal vector	2.1
	$\boldsymbol{n}_{\partial\Omega}$	Outward pointing unit normal vector with respect to Ω	4.5.1
	$\boldsymbol{n}_E$	Outward pointing unit normal vector with respect to a face E	3.3
	$\boldsymbol{n}_{\partial K}$	Outward pointing unit normal vector with respect to the boundary of a mesh cell K	3.3
	N	Norm of the convective term	6.1
	N_{div}	Norm of the convective term	6.1
	N_{div}^h	Norm of the discrete convective term	6.2
	N_L	Parameter that determines the dimension of L^H	8.8.6
	N_K	Dimension of $P(K)$	B.2
	N_p	Dimension of the pressure finite element space	4.3
	N_v	Dimension of one component of the velocity finite element space	4.3
	$N_{l,i}(\cdot)$	Global functional on multigrid level l	9.2.2
	$N_{l,i}^K(\cdot)$	Local functional on multigrid level l	9.2.2
	$\boldsymbol{N}$	Nonlinear operator	6.3
	$\boldsymbol{N}_{\text{lin}}$	Linear operator	6.3
P	p	Dimensionless pressure	2.2
	$\underline{p}$	Vector of pressure degrees of freedom	4.3
	$\hat{p}$	Function in projection methods	7.5
	p_{Bern}	Bernoulli pressure	6.1.1
	P	Dimensional pressure	2.2
	$P(K)$	Finite-dimensional space on a mesh cell K	B.2
	$P(\hat{K})$	Finite-dimensional space on the reference mesh cell $\hat{K}$	B.2
	$P_E^h\boldsymbol{v}$	Interpolation operator for Crouzeix–Raviart finite element	3.6.5
	$P_k(K)$	Finite-dimensional polynomial space on K	B.2
	P_k	Polynomial finite element space	B.2
	P_{For}^h	Fortin operator	3.5.1
	P_{Cle}^h	Clément interpolation operator	C.2
	P_{helm}	Helmholtz projection	3.7
	$P_{L^2}^h$	$L^2(\Omega)$ projection to some finite-dimensional space	C.1
	P_{loc}^h	Local projection operator	5.4
	$P_{H^1}^h$	Elliptic projection to some finite-dimensional space	C.1
	P_{l-1}^l	Prolongation operator from multigrid level $(l-1)$ to level l	9.2.2
Q	Q	Space in abstract saddle point problem	3.1
	Q	Pressure space for continuous problems	3.2
	$\tilde{Q}$	Pressure space with functions from $H^1(\Omega)$	4.5.1
	Q_k	Polynomial finite element space	B.2
	Q^h	Finite element pressure space	3.3
	Q_{std}^h	Finite element pressure space with standard basis	4.3
	Q_M^h	Finite element pressure space on macroelement	3.5.4
	$Q_{0,M}^h$	Finite element pressure space on macroelement	3.5.4
	$Q_{\text{grad},M}^h$	Finite element pressure space on macroelement	3.5.4
	$\mathcal{Q}^h$	Set of pressure degrees of freedom	9.2.2
	$\mathcal{Q}_j^h$	Set of local pressure degrees of freedom	9.2.2

R	r	Pressure of turbulence models	8.3.2
	$\underline{r}$	Residual vector	9.1
	R_k	Space of polynomials on different mesh cells	3.5.4
	R_l^{l-1}	Function restriction from multigrid level l to level $(l-1)$	9.2.2
	$(R')_l^{l-1}$	Defect restriction from multigrid level l to level $(l-1)$	9.2.2
	Re	Reynolds number	2.3
	Re_τ	Reynolds number based on friction velocity	D.3
	$R^h_{ij,\mathrm{mean}}$	Reynolds stress for finite element velocity	D.3
	$\mathbf{Res}$	Residual vector in two-scale VMS methods	8.8.1
	$\mathcal{R}, \mathcal{R}_a, \widetilde{\mathcal{R}}_a$	Remainder terms	6.4
S	S	Schur complement matrix	9.2.3
	$\mathbb{S}$	Stress tensor	2.2
	S_d	Approximation of the Schur complement matrix	9.2.3
	S_K	Barycenter of a mesh cell K	B.2
	S_l^h	Finite element space on multigrid level l	9.2.2
	St	Strouhal number	2.3
T	t	Dimensionless time	2.3
	t_n	Discrete time	7.3.1
	t_λ	Kolmogorov time scale	8.1
	$\boldsymbol{t}$	Cauchy stress vector	2.2
	$\boldsymbol{t}_k$	Tangential vector	2.4
	T	Final time for the flow problem	2.1
	$T(\cdot;\cdot,\cdot)$	Turbulence model in VMS methods	8.8
	$\mathbb{T}$	Subgrid scale tensor, Reynolds stress tensor	8.2.2
	$\mathcal{T}^h$	Triangulation	B.2
	$\mathcal{T}_l$	Triangulation on multigrid level l	9.2.2
	T^*	Characteristic time scale of the flow problem	2.3
U	$\boldsymbol{u}$	Dimensionless velocity	2.3
	$\underline{u}$	Vector of velocity degrees of freedom	4.3
	u	Solution of Primal problem	6.4
	u^h	Solution of discrete primal problem	6.4
	$\tilde{\boldsymbol{u}}$	Intermediate velocity in pressure-correction schemes	7.5
	u_λ	Kolmogorov velocity scale	8.1
	$\boldsymbol{u}^h_{\mathbf{mean}}$	Mean finite element velocity	D.3
	U	Characteristic velocity scale of the flow problem	2.3
	$\boldsymbol{U}$	Vector of velocity and pressure	8.8.1
	U_τ	Friction velocity	8.1

V	$\boldsymbol{v}$	Dimensional velocity	2.1
	$\overline{\boldsymbol{v}}_E$	Integral mean value of $\boldsymbol{v}$ on the face E	4.2.3
	V	Space in abstract saddle point problem	3.1
	V	Velocity space for continuous problems	3.2
	$V(r)$	Manifold in V	3.1
	V_0	Subspace of V	3.1
	$\tilde{V}'$	Subset of V'	3.1
	V_{div}	Space of weakly divergence-free functions	3.2
	V^h	Finite element velocity space	3.3
	V^h_{div}	Space of discretely divergence-free functions	3.3
	V^n_{div}	Finite-dimensional space of divergence-free functions	7.1
	$V^h_{\text{div,div}}$	Intersection of V_{div} and V^h_{div}	4.6.1
	$V^h_{0,M}$	Finite element velocity space on macroelement	3.5.4
	V^h_l	Finite element space on multigrid level l	9.2.2
	$\mathcal{V}^h$	Set of velocity degrees of freedom	9.2.2
	$\mathcal{V}^h_j$	Set of local velocity degrees of freedom	9.2.2
W	$\boldsymbol{w}$	Velocity field of turbulence models, approximation of the large scales	8.3.2
	$\boldsymbol{w}_{\text{drag}}$	Direction of drag force	D
	$\boldsymbol{w}_{\text{lift}}$	Direction of lift force	D
X	$\underline{x}$	Solution vector of global algebraic saddle	9
Y	$\underline{y}$	Right-hand side vector of global algebraic saddle point problem	9
	y_c	Cut-off wave number	8.2
	y^+	Wall unit	8.1
Z	z	Solution of dual linearized problem	6.4
	z^h	Solution of discrete dual linearized problem	6.4
$\boldsymbol{\beta}$	β	Friction parameter	2.4
	β_{is}	Inf-sup parameter for abstract and continuous problems	3.1
	β^h_{is}	Inf-sup parameter for discrete problems	3.3
	$\beta^h_{\text{is,sma}}$	Inf-sup parameter for discrete Smagorinsky problem	8.3.3.2
	$\beta_{\text{is,Bab}}$	Inf-sup parameter for Babuška's condition	3.1
$\boldsymbol{\gamma}$	γ^h	Bound of Fortin operator	3.5.1
	Γ	Boundary of a domain Ω	2.4
	Γ_{body}	Boundary of a body in the flow	D
	Γ_{diri}	Part of the boundary with Dirichlet conditions	2.4
	Γ_{nosl}	Part of the boundary with no–slip conditions	2.4
	Γ_{donot}	Part of the boundary with outflow conditions	2.4
	Γ_{slip}	Part of the boundary with free slip conditions	2.4
	Γ_{slfr}	Part of the boundary with slip with friction conditions	2.4

$\boldsymbol{\delta}$	δ	Filter width	8.2
	$\boldsymbol{\delta}$	Tensor of stabilization parameters	8.8.2
	δ_K, δ_E, δ	Stabilization parameters in residual based stabilizations	5.3
	δ_ν	Viscous length scale	8.1
	δ_{ij}	Kronecker symbol	
	$\delta\rho$	Remainder term	6.4
	Δt_{n+1}	Length of the time step	7.3.1
	Δ^h	Discrete Laplacian	8.6
$\boldsymbol{\varepsilon}$	ε	Rate of dissipation of turbulent energy	8.1
	$\varepsilon_{\mathrm{Ler}}$	Rate of dissipation of turbulent energy for Leray-α model	8.6
$\boldsymbol{\zeta}$	ζ	Second order viscosity of the fluid	2.2
$\boldsymbol{\theta}$	$\theta,\ \theta_1,\ \ldots,\ \theta_4$	Parameters in θ-schemes	7.3.1
$\boldsymbol{\lambda}$	λ	Size of the smallest eddies, Kolmogorov length scale	8.1
	$\lambda_{\nu,\mathrm{Ler}}$	Size of the smallest eddies for Leray model	8.6
	λ_i	Barycentric coordinate	B.2
	λ_{2d}	Kraichnan dissipation length	8.1
	λ_ν	Viscous length scale	8.1
$\boldsymbol{\mu}$	μ	Dynamic viscosity of the fluid	2.2
	μ, μ_K	Stabilization parameters for grad-div stabilization	5.3
$\boldsymbol{\nu}$	ν	Kinematic viscosity of the fluid	2.2
	ν	Dimensionless viscosity	2.3
	ν_{Sm}	Factor in turbulent viscosity	8.3.2
	ν_{T}	Turbulent viscosity	8.3.1
$\boldsymbol{\rho}$	ρ	Density	2.1
	ρ_0	Constant density	2.1
	ρ_K	Diameter of the largest ball contained in K	C.1
	$\rho(\cdot)$	Residual	6.4
	$\rho^*(\cdot)$	Dual residual	6.4
$\boldsymbol{\tau}$	τ_{w}	Wall shear stress	8.1
$\boldsymbol{\phi}$	ϕ	Stream function	3.7
	$\phi_{K,i}$	Local basis functions	B.2
	$\boldsymbol{\phi}_i^h$	Basis function for velocity finite element space	4.3
	ϕ_i^h	Scalar component of $\boldsymbol{\phi}_i^h$	4.3
	Φ	Map from solution to right-hand side in abstract linear saddle point problem	3.1
	Φ_i, $\Phi_{K,i}$	Global and local functionals	B.2
	$\Phi(v)$	Linear and continuous functionals	B.2
	$\hat{\Phi}_i$	Local functionals on reference cell $\hat{K}$	B.2
$\boldsymbol{\varphi}$	φ	Function in projection methods	7.5
$\boldsymbol{\psi}$	ψ_i^h	Basis function for pressure finite element space	4.3

$\boldsymbol{\omega}$	ω	Volume in Ω	2.1
	$\boldsymbol{\omega}$	Vorticity	8.1
	ω_{pres}	Parameter in the norm $\|(\cdot)\|_{\mathrm{spg,p}}$	5.3
	ω_E	Union of mesh cells that have the common face E	4.4
	ω_i	Union of mesh cells	B.2
	ω_K	Union of mesh cells that have a common face with cell K	4.4
	$\omega(K)$	Union of mesh cells in a neighborhood of mesh cell K	3.3
	Ω	Domain of definition for a partial differential equation	2.1
$\bullet$	∂K	Boundary of a mesh cell K	B.2
	$\partial\Omega$	Boundary of a domain Ω	2.4
	$(\cdot,\cdot)$	Inner product in $L^2(\Omega)$	3.2
	$(\cdot,\cdot)_\omega$	Inner product in $L^2(\omega)$ with $\omega \subset \Omega$	
	$\langle\cdot,\cdot\rangle_{V',V}$	Dual pairing	3.1
	$\overline{(\cdot)}^h$	Discrete differential filter	8.6
	$\|\cdot\|_2$	Euclidean norm of a vector	
	$\|\cdot\|_{\mathrm{F}}$	Frobenius norm of a matrix or a tensor	
	$\langle\cdot\rangle_s$	Average in space	D.3
	$\langle\cdot\rangle_t$	Average in time	D.3
	$[\|\cdot\|]_E$	Jump of a function across a face E	3.3
	$\{\{\cdot\}\}_E$	Average of a function on a face E	3.3

References

Abdi N (2015) Turbulence modelling of the Navier-Stokes equations using the NS-alpha approach. Master thesis, Freie Universität Berlin

Acosta G, Durán RG (1999) The maximum angle condition for mixed and nonconforming elements: application to the Stokes equations. SIAM J Numer Anal 37:18–36 (electronic)

Adams RA (1975) Sobolev spaces. Pure and applied mathematics, vol 65. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York/London, pp xviii+268

Ahmed N, Chacón Rebollo T, John V, Rubino S (2016) A review of variational multiscale methods for the simulation of turbulent incompressible flows. Arch Comput Methods Eng 1–50 (in press)

Ainsworth M, Oden JT (2000) A posteriori error estimation in finite element analysis. Pure and applied mathematics (New York). Wiley-Interscience [Wiley], New York, pp xx+240

Aldama AA (1990) Filtering techniques for turbulent flow simulation. Lecture notes in engineering, vol 56. Springer, Berlin, pp viii+397

Amann H (2000) On the strong solvability of the Navier-Stokes equations. J Math Fluid Mech **2**:16–98

Amestoy PR, Duff IS, L'Excellent J-Y, Koster J (2001) A fully asynchronous multifrontal solver using distributed dynamic scheduling. SIAM J Matrix Anal Appl 23:15–41 (electronic)

Amestoy PR, Guermouche A, L'Excellent J-Y, Pralet S (2006) Hybrid scheduling for the parallel solution of linear systems. Parallel Comput 32:136–156

Arndt D, Dallmann H, Lube G (2015) Local projection FEM stabilization for the time-dependent incompressible Navier-Stokes problem. Numer Methods Partial Differ Equ 31:1224–1250

Arnold D, Qin J (1992) Quadratic velocity/linear pressure Stokes elements. In: Vichnevetsky R, Knight D, Richter G (eds) Advances in computer methods for partial differential equations VII. IMACS, New Brunswick, pp 28–34

Arnold DN, Brezzi F, Fortin M (1984) A stable finite element for the Stokes equations. Calcolo 21:337–344

Arnold DN, Boffi D, Falk RS (2002) Approximation by quadrilateral finite elements. Math Comp 71:909–922 (electronic)

Aubin J-P (1967) Behavior of the error of the approximate solutions of boundary value problems for linear elliptic operators by Gelerkin's and finite difference methods. Ann Scuola Norm Sup Pisa (3) 21:599–637

Ayuso B, García-Archilla B, Novo J (2005). The postprocessed mixed finite-element method for the Navier-Stokes equations. SIAM J Numer Anal 43:1091–1111

Babuška I (1970/1971) Error-bounds for finite element method. Numer Math 16:322–333

V. John, *Finite Element Methods for Incompressible Flow Problems*, Springer Series in Computational Mathematics 51, DOI 10.1007/978-3-319-45750-5

Babuška I (1972/1973) The finite element method with Lagrangian multipliers. Numer Math 20:179–192

Badia S, Codina R, Gutiérrez-Santacreu JV (2010) Long-term stability estimates and existence of a global attractor in a finite element approximation of the Navier-Stokes equations with numerical subgrid scale modeling. SIAM J Numer Anal 48:1013–1037

Baker GA (1976) Galerkin approximations for the Navier-Stokes equations. Technical Report. Havard University

Balay S, Abhyankar S, Adams MF, Brown J, Brune P, Buschelman K, Dalcin L, Eijkhout V, Gropp WD, Kaushik D, Knepley MG, McInnes LC, Rupp K, Smith BF, Zampini S, Zhang H, Zhang H (2016) PETSc Web page. http://www.mcs.anl.gov/petsc

Bangerth W, Rannacher R (2003) Adaptive finite element methods for differential equations. Lectures in mathematics ETH Zürich. Birkhäuser Verlag, Basel, pp viii+207

Bangerth W, Hartmann R, Kanschat G (2007) deal.II—a general-purpose object-oriented finite element library. ACM Trans Math Softw 33. Art. 24, 27

Barbato D, Berselli LC, Grisanti CR (2007) Analytical and numerical results for the rational large eddy simulation model. J Math Fluid Mech 9:44–74

Barrenechea GR, Valentin F (2010a) Consistent local projection stabilized finite element methods. SIAM J Numer Anal 48:1801–1825

Barrenechea GR, Valentin F (2010b) A residual local projection method for the Oseen equation. Comput Methods Appl Mech Eng 199:1906–1921

Barrenechea GR, Valentin F (2011) Beyond pressure stabilization: a low-order local projection method for the Oseen equation. Int J Numer Methods Eng 86:801–815

Barth T, Bochev P, Gunzburger M, Shadid J (2004) A taxonomy of consistently stabilized finite element methods for the Stokes problem. SIAM J Sci Comput 25:1585–1607

Bause M (1997) Optimale Konvergenzraten für voll diskretisierte Navier-Stokes-Approximationen höherer Ordnung in Gebieten mit Lipschitz-Rand. PhD thesis, Univ.-GH Paderborn, FB Mathematik/ Informatik, Paderborn, pp 175

Bazilevs Y, Calo VM, Cottrell JA, Hughes TJR, Reali A, Scovazzi G (2007) Variational multiscale residual-based turbulence modeling for large eddy simulation of incompressible flows. Comput Methods Appl Mech Eng 197:173–201

Becker R (2000) An optimal-control approach to a posteriori error estimation for finite element discretizations of the Navier-Stokes equations. East-West J Numer Math 8:257–274

Becker R, Braack M (2001) A finite element pressure gradient stabilization for the Stokes equations based on local projections. Calcolo 38:173–199

Becker R, Rannacher R (1996) A feed-back approach to error control in finite element methods: basic analysis and examples. East-West J Numer Math 4:237–264

Becker R, Rannacher R (1998) Weighted a posteriori error control in FE methods. In: Bock HG, Kanschat G, Rannacher R, Brezzi F, Glowinski R, Kuznetsov YA, Périaux J (eds) ENUMATH 97. Proceedings of the 2nd European conference on numerical mathematics and advanced applications held at the University of Heidelberg, Heidelberg, September 28–October 3, 1997. World Scientific Publishing Co., Inc., River Edge, NJ, pp 621–637

Becker R, Rannacher R (2001) An optimal control approach to a posteriori error estimation in finite element methods. Acta Numer 10:1–102

Benzi M, Olshanskii MA (2006) An augmented Lagrangian-based approach to the Oseen problem. SIAM J Sci Comput 28:2095–2113

Benzi M, Wang Z (2011) Analysis of augmented Lagrangian-based preconditioners for the steady incompressible Navier-Stokes equations. SIAM J Sci Comput 33:2761–2784

Benzi M, Golub GH, Liesen J (2005) Numerical solution of saddle point problems. Acta Numer 14:1–137

Bercovier M, Pironneau O (1979) Error estimates for finite element method solution of the Stokes problem in the primitive variables. Numer Math 33:211–224

Bernardi C, Raugel G (1985) Analysis of some finite elements for the Stokes problem. Math Comp 44:71–79

Bernardi C, Chacón Rebollo T, Yakoubi D (2015) Finite element discretization of the Stokes and Navier-Stokes equations with boundary conditions on the pressure. SIAM J Numer Anal 53:1256–1279

Berrone S, Marro M (2009) Space-time adaptive simulations for unsteady Navier-Stokes problems. Comput Fluids 38:1132–1144

Berselli LC, Iliescu T (2003) A higher-order subfilter-scale model for large eddy simulation. J Comput Appl Math 159:411–430

Berselli LC, John V (2006) Asymptotic behaviour of commutation errors and the divergence of the Reynolds stress tensor near the wall in the turbulent channel flow. Math Methods Appl Sci 29:1709–1719

Berselli LC, Lewandowski R (2012) Convergence of approximate deconvolution models to the mean Navier-Stokes equations. Ann Inst H Poincaré Anal Non Linéaire 29:171–198

Berselli LC, Galdi GP, Iliescu T, Layton WJ (2002) Mathematical analysis for the rational large eddy simulation model. Math Models Methods Appl Sci 12:1131–1152

Berselli LC, Iliescu T, Layton WJ (2006) Mathematics of large eddy simulation of turbulent flows. Scientific computation. Springer, Berlin, pp xviii+348

Berselli LC, Grisanti CR, John V (2007) Analysis of commutation errors for functions with low regularity. J Comput Appl Math 206:1027–1045

Besier M, Rannacher R (2012) Goal-oriented space-time adaptivity in the finite element Galerkin method for the computation of nonstationary incompressible flow. Int J Numer Methods Fluids 70:1139–1166

Besier M, Wollner W (2012) On the pressure approximation in nonstationary incompressible flow simulations on dynamically varying spatial meshes. Int J Numer Methods Fluids 69:1045–1064

Bochev P, Gunzburger M (2004) An absolutely stable pressure-Poisson stabilized finite element method for the Stokes equations. SIAM J Numer Anal 42:1189–1207

Bochev PB, Dohrmann CR, Gunzburger MD (2006) Stabilization of low-order mixed finite elements for the Stokes equations. SIAM J Numer Anal 44:82–101 (electronic)

Boffi D (1994) Stability of higher order triangular Hood-Taylor methods for the stationary Stokes equations. Math Models Methods Appl Sci 4:223–235

Boffi D (1997) Three-dimensional finite element methods for the Stokes problem. SIAM J Numer Anal 34:664–670

Boffi D, Brezzi F, Fortin M (2008) Finite elements for the Stokes problem. In: Boffi D , Gastaldi L (eds) Mixed finite elements, compatibility conditions, and applications. Lecture notes in mathematics, vol 1939. Lectures given at the C.I.M.E. Summer School held in Cetraro, June 26–July 1, 2006. Springer, Berlin, pp 45–100

Boffi D, Brezzi F, Fortin M (2013) Mixed finite element methods and applications. Springer series in computational mathematics, vol 44. Springer, Heidelberg, pp xiv+685

Boland JM, Nicolaides RA (1983) Stability of finite elements under divergence constraints. SIAM J Numer Anal 20:722–731

Boland JM, Nicolaides RA (1985) Stable and semistable low order finite elements for viscous flows. SIAM J Numer Anal 22:474–492

Boussinesq J (1877) Essai sur la théorie des eaux courantes. Mémoires, L'Académie des Sciences de l'Institut de France, vol XXIII. Imprimerie Nationale, Paris

Bowers AL, Le Borne S, Rebholz LG (2014) Error analysis and iterative solvers for Navier-Stokes projection methods with standard and sparse grad-div stabilization. Comput Methods Appl Mech Eng 275:1–19

Braack M, Burman E (2006) Local projection stabilization for the Oseen problem and its interpretation as a variational multiscale method. SIAM J Numer Anal 43:2544–2566 (electronic)

Braack M, Lube G (2009) Finite elements with local projection stabilization for incompressible flow problems. J Comput Math 27:116–147

Braack M, Mucha PB (2014) Directional do-nothing condition for the Navier-Stokes equations. J Comput Math 32:507–521

Braack M, Burman E, John V, Lube G (2007) Stabilized finite element methods for the generalized Oseen problem. Comput Methods Appl Mech Eng 196:853–866

Braess D, Blömer C (1990) A multigrid method for a parameter dependent problem in solid mechanics. Numer Math 57:747–761

Brennecke C, Linke A, Merdon C, Schöberl J (2015) Optimal and pressure-independent L^2 velocity error estimates for a modified Crouzeix-Raviart Stokes element with BDM reconstructions. J Comput Math 33:191–208

Brenner SC (2003) Poincaré-Friedrichs inequalities for piecewise H^1 functions. SIAM J Numer Anal 41:306–324

Brenner SC, Scott LR (2008) The mathematical theory of finite element methods. Texts in applied mathematics, vol 15, 3rd edn. Springer, New York, pp xviii+397

Breuer M (1998) Large eddy simulation of the subcritical flow past a circular cylinder: numerical and modeling aspects. Int J Numer Methods Fluids 28:1281–1302

Brezzi F (1974) On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. Rev Française Automat Informat Recherche Opérationnelle Sér. Rouge 8:129–151

Brezzi F, Bathe K-J (1990) A discourse on the stability conditions for mixed finite element formulations. Comput Methods Appl Mech Eng 82:27–57. Reliability in computational mechanics (Austin, TX, 1989)

Brezzi F, Douglas J Jr (1988) Stabilized mixed methods for the Stokes problem. Numer Math 53:225–235

Brezzi F, Falk RS (1991) Stability of higher-order Hood-Taylor methods. SIAM J Numer Anal 28:581–590

Brezzi F, Fortin M (2001) A minimal stabilisation procedure for mixed finite element methods. Numer Math 89:457–491

Brezzi F, Pitkäranta J (1984) On the stabilization of finite element approximations of the Stokes equations. In: Efficient solutions of elliptic systems (Kiel, 1984). Notes Numerical Fluid Mechanics, vol 10. Friedr. Vieweg, Braunschweig, pp 11–19

Brezzi F, Russo A (1994) Choosing bubbles for advection-diffusion problems. Math Models Methods Appl Sci 4:571–587

Briggs WL, Henson VE, McCormick SF (2000) A multigrid tutorial, 2nd edn. Society for industrial and applied mathematics (SIAM), Philadelphia, PA, pp xii+193

Bristeau MO, Glowinski R, Périaux J (1987) Numerical methods for the Navier-Stokes equations. Applications to the simulation of compressible and incompressible viscous flows. In: Finite elements in physics (Lausanne, 1986). North-Holland, Amsterdam, pp 73–187

Brooks AN, Hughes TJR (1982) Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. Comput Methods Appl Mech Eng 32:199–259. FENOMECH ’81, part I (Stuttgart, 1981)

Bruneau C-H, Saad M (2006) The 2D lid-driven cavity problem revisited. Comput Fluids 35:326–348

Burman E, Fernández MA (2007) Continuous interior penalty finite element method for the time-dependent Navier-Stokes equations: space discretization and convergence. Numer Math 107:39–77

Burman E, Fernández MA (2008) Galerkin finite element methods with symmetric pressure stabilization for the transient Stokes equations: stability and convergence analysis. SIAM J Numer Anal 47:409–439

Burman E, Fernández MA (2011) Analysis of the PSPG method for the transient Stokes’ problem. Comput Methods Appl Mech Eng 200:2882–2890

Burman E, Hansbo P (2004) Edge stabilization for Galerkin approximations of convection-diffusion-reaction problems. Comput Methods Appl Mech Eng 193:1437–1453

Burman E, Hansbo P (2006) A stabilized non-conforming finite element method for incompressible flow. Comput Methods Appl Mech Eng 195:2881–2899

Burman E, Fernández MA, Hansbo P (2006) Continuous interior penalty finite element method for Oseen’s equations. SIAM J Numer Anal 44:1248–1274

Burton GC, Dahm WJA (2005a) Multifractal subgrid-scale modeling for large-eddy simulation. I. Model development and a priori testing. Phys Fluids 17:075111, 16

Burton GC, Dahm WJA (2005b) Multifractal subgrid-scale modeling for large-eddy simulation. II. Backscatter limiting and a posteriori evaluation. Phys Fluids 17:075112, 19

Bychenkov YV, Chizonkov EV (1999) Optimization of one three-parameter method of solving an algebraic system of the Stokes type. Russ J Numer Anal Math Modell 14:429–440

Caiazzo A, Iliescu T, John V, Schyschlowa S (2014) A numerical investigation of velocity-pressure reduced order models for incompressible flows. J Comput Phys 259:598–616

Camassa R, Holm DD (1993) An integrable shallow water equation with peaked solitons. Phys Rev Lett 71:1661–1664

Carathéodory C (1918) Vorlesungen ber reelle Funktionen. Teubner, Leipzig und Berlin. Reprint of 3rd edition. Chelsea Publishing Co., New York (1968)

Carroll R, Duff G, Friberg J, Gobert J, Grisvard P, Nečas J, Seeley R (1966) Équations aux dérivées partielles. Séminaire de Mathématiques Supérieures. 19. Montréal: Les Presses de l'Université de Montréal. 142 p

Case MA, Ervin VJ, Linke A, Rebholz LG (2011) A connection between Scott-Vogelius and grad-div stabilized Taylor-Hood FE approximations of the Navier-Stokes equations. SIAM J Numer Anal 49:1461–1481

Cazemier W, Verstappen RWCP, Veldman AEP (1998) Proper orthogonal decomposition and low-dimensional models for driven cavity flows. Phys Fluids 10:1685–1699

Chacón Rebollo T, Lewandowski R (2014) Mathematical and numerical foundations of turbulence models and applications. Modeling and simulation in science, engineering and technology. Birkhäuser/Springer, New York, pp xviii+517

Chacón Rebollo T, Gómez Mármol M, Girault V, Sánchez Muñoz I (2013) A high order term-by-term stabilization solver for incompressible flow problems. IMA J Numer Anal 33:974–1007

Chen S, Foias C, Holm DD, Olson E, Titi ES, Wynne S (1998) Camassa-Holm equations as a closure model for turbulent channel and pipe flow. Phys Rev Lett 81:5338–5341

Chen S, Foias C, Holm DD, Olson E, Titi ES, Wynne S (1999a) The Camassa-Holm equations and turbulence. Phys D 133:49–65. Predictability: quantifying uncertainty in models of complex phenomena (Los Alamos, NM, 1998)

Chen S, Foias C, Holm DD, Olson E, Titi ES, Wynne S (1999b) A connection between the Camassa-Holm equations and turbulent flows in channels and pipes. Phys Fluids 11:2343–2353. The international conference on turbulence (Los Alamos, NM, 1998)

Cheskidov A, Holm DD, Olson E, Titi ES (2005) On a Leray-α model of turbulence. Proc R Soc Lond Ser A Math Phys Eng Sci 461:629–649

Chizhonkov EV, Olshanskii MA (2000) On the domain geometry dependence of the LBB condition. M2AN Math Model Numer Anal 34:935–951

Choi H, Moin P (1994) Effects of the computational time step on numerical solutions of turbulent flow. J Comput Phys 113:1–4

Chorin AJ (1968) Numerical solution of the Navier-Stokes equations. Math Comp 22:745–762

Ciarlet PG (1978) The finite element method for elliptic problems. Studies in mathematics and its applications, vol 4. North-Holland Publishing Co., Amsterdam , pp xix+530

Ciarlet P Jr, Huang J, Zou J (2003) Some observations on generalized saddle-point problems. SIAM J Matrix Anal Appl 25:224–236

Clark RA, Ferziger JH, Reynolds WC (1979) Evaluation of subgrid-scale models using an accurately simulated turbulent flow. J Fluid Mech 91:1–16

Clément P (1975) Approximation by finite element functions using local regularization. Rev. Française Automat. Informat. Recherche Opérationnelle Sér. RAIRO Analyse Numérique 9:77–84

Clough R, Tocher J (1965) Finite element stiffness matrices for analysis of plates in bending. In: Proceedings of the conference on matrix methods in structural mechanics. Wright Patterson A.F.B., OH, pp 515–545

Codina R (2001a) Pressure stability in fractional step finite element methods for incompressible flows. J Comput Phys 170:112–140

Codina R (2001b) A stabilized finite element method for generalized stationary incompressible flows. Comput Methods Appl Mech Eng 190:2681–2706

Codina R (2002) Stabilized finite element approximation of transient incompressible flows using orthogonal subscales. Comput Methods Appl Mech Eng 191:4295–4321

Codina R (2008) Analysis of a stabilized finite element approximation of the Oseen equations using orthogonal subscales. Appl Numer Math 58:264–283

Codina R, Blasco J (1997) A finite element formulation for the Stokes problem allowing equal velocity-pressure interpolation. Comput Methods Appl Mech Eng 143:373–391

Codina R, Blasco J (2000) Analysis of a pressure-stabilized finite element approximation of the stationary Navier-Stokes equations. Numer Math 87:59–81

Codina R, Principe J, Guasch O, Badia S (2007) Time dependent subscales in the stabilized finite element approximation of incompressible flow problems. Comput Methods Appl Mech Eng 196:2413–2430

Codina R, Principe J, Badia S (2011) Dissipative structure and long term behavior of a finite element approximation of incompressible flows with numerical subgrid scale modeling. Multiscale methods in computational mechanics. Lecture notes in applied and computational mechanics, vol 55. Springer, Dordrecht, pp 75–93

Coletti P (1997) A global existence theorem for large eddy simulation turbulence model. Math Models Methods Appl Sci 7:579–591

Colomés O, Badia S, Codina R, Principe J (2015) Assessment of variational multiscale models for the large eddy simulation of turbulent incompressible flows. Comput Methods Appl Mech Eng 285:32–63

Connors J (2010) Convergence analysis and computational testing of the finite element discretization of the Navier-Stokes alpha model. Numer Methods Partial Differ Equ 26:1328–1350

Constantin P, Foias C (1988) Navier-Stokes equations. Chicago lectures in mathematics. University of Chicago Press, Chicago, IL, pp x+190

Constantin P, Foias C, Temam R (1985a) Attractors representing turbulent flows. Mem Am Math Soc 53:vii+67

Constantin P, Foias C, Manley OP, Temam R (1985b) Determining modes and fractal dimension of turbulent flows. J Fluid Mech 150:427–440

Cools R, Rabinowitz P (1993) Monomial cubature rules since "Stroud": a compilation. J Comput Appl Math 48:309–326

Crouzeix M, Raviart P-A (1973) Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I. Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge 7:33–75

Cuff VM, Dunca AA, Manica CC, Rebholz LG (2015) The reduced order NS-α model for incompressible flow: theory, numerical analysis and benchmark testing. ESAIM Math Model Numer Anal 49:641–662

Dahl O, Wille S (1992) An ILU preconditioner with coupled node fill-in for iterative solution of the mixed finite element formulation of the 2D and 3D Navier-Stokes equations. Int J Numer Methods Fluids 15:525–544

Dari E, Durán R, Padra C (1995) Error estimators for nonconforming finite element approximations of the Stokes problem. Math Comp 64:1017–1033

Dauge M (1989) Stationary Stokes and Navier-Stokes systems on two- or three-dimensional domains with corners. I. Linearized equations. SIAM J Math Anal 20:74–97

Davidson PA (2004) Turbulence: an introduction for scientists and engineers. Oxford University Press, Oxford, pp xx+657

Davis TA (2004) Algorithm 832: UMFPACK V4.3—an unsymmetric-pattern multifrontal method. ACM Trans Math Softw 30:196–199

de Frutos J, García-Archilla B, Novo J (2008) The postprocessed mixed finite-element method for the Navier-Stokes equations: refined error bounds. SIAM J Numer Anal 46:201–230

de Frutos J, García-Archilla B, John V, Novo J (2016a) Analysis of the grad-div stabilization for the time-dependent Navier–Stokes equations with inf-sup stable finite elements. (submitted) de Frutos J, García-Archilla B, John V, Novo J (2016b) Grad-div stabilization for the evolutionary Oseen problem with inf-sup stable finite elements. J Sci Comput 66:991–1024

de Frutos J, John V, Novo J (2016c) Projection methods for incompressible flow problems with WENO finite difference schemes. J Comput Phys 309:368–386

Demkowicz L (2006) Babuška ↔ Brezzi ?? ICES Report 06-08, The University of Texas at Austin, Institute for Computational Engineering and Sciences

Dohrmann CR, Bochev PB (2004) A stabilized finite element method for the Stokes problem based on polynomial pressure projections. Int J Numer Methods Fluids 46:183–201

Douglas J Jr, Dupont T (1976) Interior penalty procedures for elliptic and parabolic Galerkin methods. In: Computing methods in applied sciences (second international symposium, Versailles, 1975). Lecture Notes in Physics, vol 58 Springer, Berlin, pp 207–216

Douglas J Jr, Wang JP (1989) An absolutely stabilized finite element method for the Stokes problem. Math Comp 52:495–508

Du Q, Gunzburger MD (1990) Finite-element approximations of a Ladyzhenskaya model for stationary incompressible viscous flow. SIAM J Numer Anal 27:1–19

Du Q, Gunzburger MD (1991) Analysis of a Ladyzhenskaya model for incompressible viscous flow. J Math Anal Appl 155:21–45

Dunca A, Epshteyn Y (2006) On the Stolz-Adams deconvolution model for the large-eddy simulation of turbulent flows. SIAM J Math Anal 37:1890–1902 (electronic)

Dunca A, John V, Layton WJ (2004) The commutation error of the space averaged Navier-Stokes equations on a bounded domain. In: Contributions to current challenges in mathematical fluid mechanics. Advances in Mathematical Fluid Mechanics. Birkhäuser, Basel, pp 53–78

Duvaut G, Lions J-L (1972) Les inéquations en mécanique et en physique. Travaux et Recherches Mathématiques, vol 21. Dunod, Paris, pp xx+387

Elman HC, Tuminaro RS (2009) Boundary conditions in approximate commutator preconditioners for the Navier-Stokes equations. Electron Trans Numer Anal 35:257–280

Elman H, Howle VE, Shadid J, Shuttleworth R, Tuminaro R (2006) Block preconditioners based on approximate commutators. SIAM J Sci Comput 27:1651–1668

Elman H, Howle VE, Shadid J, Silvester D, Tuminaro R (2008a) Least squares preconditioners for stabilized discretizations of the Navier-Stokes equations. SIAM J Sci Comput 30:290–311

Elman H, Howle VE, Shadid J, Shuttleworth R, Tuminaro R (2008b) A taxonomy and comparison of parallel block multi-level preconditioners for the incompressible Navier-Stokes equations. J Comput Phys 227:1790–1808

Elman HC, Silvester DJ, Wathen AJ (2014) Finite elements and fast iterative solvers: with applications in incompressible fluid dynamics. Numerical mathematics and scientific computation, 2nd edn. Oxford University Press, Oxford, pp xiv+479

Emmrich E (1999) Discrete versions of Gronwall's lemma and their application to the numerical analysis of parabolic problems. Preprint 637. Technische Universität Berlin, Fachbereich Mathematik

Emmrich E (2001) Analysis von Zeitdiskretisierungen des inkompressiblen Navier-Stokes-Problems. Diss. Cuvillier Verlag, Göttingen; TU Berlin, Fakultät II–Mathematik und Naturwissenschaften, Berlin, pp viii+206

Emmrich E (2004a) Error of the two-step BDF for the incompressible Navier-Stokes problem. M2AN Math Model Numer Anal 38:757–764

Emmrich E (2004b) Stability and convergence of the two-step BDF for the incompressible Navier-Stokes problem. Int J Nonlinear Sci Numer Simul 5:199–209

Ern A, Guermond J-L (2004) Theory and practice of finite elements. Applied mathematical sciences, vol 159. Springer, New York, pp xiv+524

Ethier C, Steinman D (1994) Exact fully 3D Navier-Stokes solutions for benchmarking. Int J Numer Methods Fluids 19:369–375

Evans LC (2010) Partial differential equations. Graduate studies in mathematics, vol 19, 2nd edn. American Mathematical Society, Providence, RI, pp xxii+749

Falk RS, Neilan M (2013) Stokes complexes and the construction of stable finite elements with pointwise mass conservation. SIAM J Numer Anal 51:1308–1326

Farwig R (2014) On regularity of weak solutions to the instationary Navier-Stokes system: a review on recent results. Ann Univ Ferrara Sez VII Sci Mat 60:91–122

Fefferman C (2000) Existence and smoothness of the Navier-Stokes equations. http://www.claymath.org/sites/default/files/navierstokes.pdf

Ferziger JH, Perić M (1999) Computational methods for fluid dynamics, revised edn. Springer, Berlin, pp xiv+389

Filippov AF (1988) Differential equations with discontinuous righthand sides. Mathematics and its applications (Soviet Series), vol 18. Kluwer Academic Publishers Group, Dordrecht, pp x+304. Translated from the Russian

Foiaş C, Temam R (1979) Some analytic and geometric properties of the solutions of the evolution Navier-Stokes equations. J Math Pures Appl (9) 58:339–368

Foias C, Manley O, Rosa R, Temam R (2001) Navier-Stokes equations and turbulence. Encyclopedia of mathematics and its applications, vol 83. Cambridge University Press, Cambridge, pp xiv+347

Foias C, Holm DD, Titi ES (2002) The three dimensional viscous Camassa-Holm equations, and their relation to the Navier-Stokes equations and turbulence theory. J Dynam Differ Equ 14:1–35

Folland GB (1995) Introduction to partial differential equations, 2nd edn. Princeton University Press, Princeton, NJ, pp xii+324

Fortin M (1977) An analysis of the convergence of mixed finite element methods. RAIRO Anal Numér 11:341–354, iii

Franca LP, Frey SL (1992) Stabilized finite element methods. II. The incompressible Navier-Stokes equations. Comput Methods Appl Mech Eng 99:209–233

Franca LP, Hughes TJR (1988) Two classes of mixed finite element methods. Comput Methods Appl Mech Eng 69:89–129

Franca L, Nesliturk A (2001) On a two-level finite element method for the incompressible Navier-Stokes equations. Int J Numer Methods Eng 52:433–453

Franca LP, Hughes TJR, Stenberg R (1993) Stabilized finite element methods. In: Incompressible computational fluid dynamics: trends and advances. Cambridge University Press, Cambridge, pp 87–107

Friedrichs KO (1947) On the boundary-value problems of the theory of elasticity and Korn's inequality. Ann Math (2) 48:441–471

Fuhrmann J, Linke A, Langmach H, Baltruschat H (2009) Numerical calculation of the limiting current for a cylindrical thin layer flow cell. Electrochim Acta 55:430–438

Fuhrmann J, Linke A, Langmach H (2011) A numerical method for mass conservative coupling between fluid flow and solute transport. Appl Numer Math 61:530–553

Galdi GP (1994) An introduction to the mathematical theory of the Navier-Stokes equations, vol. I. Springer tracts in natural philosophy, vol 38. Springer, New York, pp xii+450. Linearized steady problems

Galdi GP (2000) An introduction to the Navier-Stokes initial-boundary value problem. In: Fundamental directions in mathematical fluid mechanics. Advances in Mathematical Fluid Mechanics. Birkhäuser, Basel, pp 1–70

Galdi GP (2011) An introduction to the mathematical theory of the Navier-Stokes equations. Steady-state problems. Springer monographs in mathematics, 2nd edn. Springer, New York, pp xiv+1018

Galdi GP, Layton WJ (2000) Approximation of the larger eddies in fluid motions. II. A model for space-filtered flow. Math Models Methods Appl Sci 10:343–350

Galvin KJ, Linke A, Rebholz LG, Wilson NE (2012) Stabilizing poor mass conservation in incompressible flow problems with large irrotational forcing and application to thermal convection. Comput Methods Appl Mech Eng 237/240:166–176

Galvin KJ, Rebholz LG, Trenchea C (2014) Efficient, unconditionally stable, and optimally accurate FE algorithms for approximate deconvolution models. SIAM J Numer Anal 52:678–707

Gamnitzer P, Gravemeier V, Wall WA (2010) Time-dependent subgrid scales in residual-based large eddy simulation of turbulent channel flow. Comput Methods Appl Mech Eng 199:819–827

Gelhard T, Lube G, Olshanskii MA, Starcke J-H (2005) Stabilized finite element schemes with LBB-stable elements for incompressible flows. J Comput Appl Math 177:243–267

Germano M, Piomelli U, Moin P, Cabot WH (1991) A dynamic subgrid-scale eddy viscosity model. Phys Fluids A 3:1760–1765

Geurts BJ, Holm DD (2003) Regularization modeling for large-eddy simulation. Phys Fluids **15**:L13–L16

Geurts BJ, Holm DD (2006) Leray and LANS-α modelling of turbulent mixing. J Turbul 7, Paper 10, 33 pp (electronic)

Ghia U, Ghia K, Shin C (1982) High-re solutions for incompressible flow using the Navier-Stokes equations and a multigrid method. J Comput Phys 48:387–411

Gilbarg D, Trudinger NS (1983) Elliptic partial differential equations of second order. In: Grundlehren der Mathematischen Wissenschaften [Fundamental principles of mathematical sciences], vol 224, 2nd edn. Springer, Berlin, pp xiii+513

Girault V, Raviart P-A (1979) Finite element approximation of the Navier-Stokes equations. Lecture notes in mathematics, vol 749. Springer, Berlin/New York, pp vii+200

Girault V, Raviart P-A (1986) Finite element methods for Navier-Stokes equations. Theory and algorithms. Springer series in computational mathematics, vol 5. Springer, Berlin, pp x+374

Girault V, Scott LR (2003) A quasi-local interpolation operator preserving the discrete divergence. Calcolo 40:1–19

Glowinski R, Le Tallec P (1989) Augmented Lagrangian and operator-splitting methods in nonlinear mechanics. SIAM studies in applied mathematics, vol 9. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, pp x+295

Gravemeier V (2006a) A consistent dynamic localization model for large eddy simulation of turbulent flows based on a variational formulation. J Comput Phys 218:677–701

Gravemeier V (2006b) Scale-separating operators for variational multiscale large eddy simulation of turbulent flows. J Comput Phys 212:400–435

Gravemeier V (2006c) The variational multiscale method for laminar and turbulent flow. Arch Comput Methods Eng 13:249–324

Gravemeier V, Wall WA, Ramm E (2004) A three-level finite element method for the instationary incompressible Navier-Stokes equations. Comput Methods Appl Mech Eng 193:1323–1366

Gravemeier V, Wall WA, Ramm E (2005) Large eddy simulation of turbulent incompressible flows by a three-level finite element method. Int J Numer Methods Fluids 48:1067–1099

Gravemeier V, Gee MW, Wall WA (2009) An algebraic variational multiscale-multigrid method based on plain aggregation for convection-diffusion problems. Comput Methods Appl Mech Eng 198:3821–3835

Gravemeier V, Gee MW, Kronbichler M, Wall WA (2010) An algebraic variational multiscale-multigrid method for large eddy simulation of turbulent flow. Comput Methods Appl Mech Eng 199:853–864

Gravemeier V, Kronbichler M, Gee MW, Wall WA (2011) An algebraic variational multiscale-multigrid method for large-eddy simulation: generalized-α time integration, Fourier analysis and application to turbulent flow past a square-section cylinder. Comput Mech 47:217–233

Gresho P, Sani R (2000) Incompressible flow and the finite element method. Vol 1: advection-diffusion. Vol 2: isothermal laminar flow. Wley, Chichester, p 1112

Guasch O, Codina R (2013) Statistical behavior of the orthogonal subgrid scale stabilization terms in the finite element large eddy simulation of turbulent flows. Comput Methods Appl Mech Eng 261/262:154–166

Guermond J-L (1999a) Stabilization of Galerkin approximations of transport equations by subgrid modeling. M2AN Math Model Numer Anal 33:1293–1316

Guermond J-L (1999b) Un résultat de convergence d'ordre deux en temps pour l'approximation des équations de Navier-Stokes par une technique de projection incrémentale. M2AN Math Model Numer Anal 33:169–189

Guermond J-L, Quartapelle L (1998) On stability and convergence of projection methods based on pressure Poisson equation. Int J Numer Methods Fluids 26:1039–1053

Guermond J L, Shen J (2004) On the error estimates for the rotational pressure-correction projection methods. Math Comp 73:1719–1737 (electronic)

Guermond JL, Minev P, Shen J (2006) An overview of projection methods for incompressible flows. Comput Methods Appl Mech Eng 195:6011–6045

Gunzburger MD (2002) The inf-sup condition in mixed finite element methods with application to the Stokes system. In: Collected lectures on the preservation of stability under discretization (Fort Collins, CO, 2001). SIAM, Philadelphia, PA, pp 93–121

Guzmán J, Neilan M (2014) Conforming and divergence-free Stokes elements on general triangular meshes. Math Comp 83:15–36

Guzmán J, Salgado AJ, Sayas F-J (2013) A note on the Ladyženskaja-Babuška-Brezzi condition. J Sci Comput 56:219–229

Hackbusch W (1985) Multigrid methods and applications. Springer series in computational mathematics, vol 4. Springer, Berlin, pp xiv+377

Hairer E, Wanner G (2010) Solving ordinary differential equations. II. Stiff and differential-algebraic problems. Springer series in computational mathematics, vol 14. Springer, Berlin, pp xvi+614. Second revised edition, paperback

Hairer E, Nørsett SP, Wanner G (1993) Solving ordinary differential equations. I. Nonstiff problems. Springer series in computational mathematics, vol 8, 2nd edn. Springer, Berlin, pp xvi+528

Hansbo P, Szepessy A (1990) A velocity-pressure streamline diffusion finite element method for the incompressible Navier-Stokes equations. Comput Methods Appl Mech Eng 84:175–192

Heister T, Rapin G (2013) Efficient augmented Lagrangian-type preconditioning for the Oseen problem using grad-div stabilization. Int J Numer Methods Fluids 71:118–134

Heywood JG, Rannacher R (1982) Finite element approximation of the nonstationary Navier-Stokes problem. I. Regularity of solutions and second-order error estimates for spatial discretization. SIAM J Numer Anal 19:275–311

Heywood JG, Rannacher R (1988) Finite element approximation of the nonstationary Navier-Stokes problem. III. Smoothing property and higher order error estimates for spatial discretization. SIAM J Numer Anal 25:489–512

Heywood JG, Rannacher R (1990) Finite-element approximation of the nonstationary Navier-Stokes problem. IV. Error analysis for second-order time discretization. SIAM J Numer Anal 27:353–384

Heywood JG, Rannacher R, Turek S (1996) Artificial boundaries and flux and pressure conditions for the incompressible Navier-Stokes equations. Int J Numer Methods Fluids 22:325–352

Hoffman J (2005) Computation of mean drag for bluff body problems using adaptive DNS/LES. SIAM J Sci Comput 27:184–207 (electronic)

Hoffman J, Johnson C (2006) A new approach to computational turbulence modeling. Comput Methods Appl Mech Eng 195:2865–2880

Hood P, Taylor C (1974) Navier–Stokes equations using mixed interpolation. In: Oden JT, Gallagher RH, Zienkiewicz OC, Taylor C (eds) Finite element methods in flow problems Huntsville Press, University of Alabama, pp 121–132

Hopf E (1951) Über die Anfangswertaufgabe für die hydrodynamischen Grundgleichungen. Math Nachr 4:213–231

Horgan CO (1995) Korn's inequalities and their applications in continuum mechanics. SIAM Rev 37:491–511

Hörmander L (1990) The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis. Springer study edition, 2nd edn. Springer, Berlin, pp xii+440

Hsu M-C, Bazilevs Y, Calo VM, Tezduyar TE, Hughes T JR (2010) Improving stability of stabilized and multiscale formulations in flow simulations at small time steps. Comput Methods Appl Mech Eng 199:828–840

Hughes TJR (1995) Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. Comput Methods Appl Mech Eng 127:387–401
Hughes TJR, Brooks A (1979) A multidimensional upwind scheme with no crosswind diffusion. Finite element methods for convection dominated flows (Papers, winter annual meeting american society of mechanical engineers, New York, 1979). AMD, vol 34. American Society of Mechanical Engineers (ASME), New York, pp 19–35
Hughes TJR, Franca LP (1987) 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. Comput Methods Appl Mech Eng 65:85–96
Hughes TJR, Franca LP, Balestra M (1986) 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. Comput Methods Appl Mech Eng 59:85–99
Hughes TJ, Mazzei L, Jansen KE (2000) Large eddy simulation and the variational multiscale method. Comput Vis Sci 3:47–59
Hunter J (2007) Matplotlib: a 2D graphics environment. Comput Sci Eng 9:90–95
Iliescu T, Layton WJ (1998) Approximating the larger eddies in fluid motion. III. The Boussinesq model for turbulent fluctuations. An Ştiinţ Univ Al I Cuza Iaşi Mat (N.S.) 44:245–261. Dedicated to Professor C. Corduneanu on the occasion of his 70th birthday.
Iliescu T, John V, Layton WJ (2002) Convergence of finite element approximations of large eddy motion. Numer Methods Partial Differ Equ 18:689–710
Iliescu T, John V, Layton WJ, Matthies G, Tobiska L (2003) A numerical study of a class of LES models. Int J Comput Fluid Dyn 17:75–85
Ingram R (2013a) A new linearly extrapolated Crank-Nicolson time-stepping scheme for the Navier-Stokes equations. Math Comp 82:1953–1973
Ingram R (2013b) Unconditional convergence of high-order extrapolations of the Crank-Nicolson, finite element method for the Navier-Stokes equations. Int J Numer Anal Model 10:257–297
Jenkins EW, John V, Linke A, Rebholz LG (2014) On the parameter choice in grad-div stabilization for the Stokes equations. Adv Comput Math 40:491–516
John V (2000) A numerical study of a posteriori error estimators for convection-diffusion equations. Comput Methods Appl Mech Eng 190:757–781
John V (2002) Higher order finite element methods and multigrid solvers in a benchmark problem for the 3D Navier-Stokes equations. Int J Numer Methods Fluids 40:775–798
John V (2004) Large eddy simulation of turbulent incompressible flows. Analytical and numerical results for a class of LES models. Lecture notes in computational science and engineering, vol 34. Springer, Berlin, pp xii+261
John V (2006) On the efficiency of linearization schemes and coupled multigrid methods in the simulation of a 3D flow around a cylinder. Int J Numer Methods Fluids 50:845–862
John V, Kaya S (2005) A finite element variational multiscale method for the Navier-Stokes equations. SIAM J Sci Comput 26:1485–1503 (electronic)
John V, Kaya S (2008) Finite element error analysis of a variational multiscale method for the Navier-Stokes equations. Adv Comput Math 28:43–61
John V, Kindl A (2008) Variants of projection-based finite element variational multiscale methods for the simulation of turbulent flows. Int J Numer Methods Fluids 56:1321–1328
John V, Kindl A (2010a) Numerical studies of finite element variational multiscale methods for turbulent flow simulations. Comput Methods Appl Mech Eng 199:841–852
John V, Kindl A (2010b) A variational multiscale method for turbulent flow simulation with adaptive large scale space. J Comput Phys 229:301–312
John V, Layton WJ (2002) Analysis of numerical errors in large eddy simulation. SIAM J Numer Anal 40:995–1020
John V, Matthies G (2001) Higher-order finite element discretizations in a benchmark problem for incompressible flows. Int J Numer Methods Fluids 37:885–903

John V, Matthies G (2004) MooNMD—a program package based on mapped finite element methods. Comput Vis Sci 6:163–169
John V, Novo J (2011) Error analysis of the SUPG finite element discretization of evolutionary convection-diffusion-reaction equations. SIAM J Numer Anal 49:1149–1176
John V, Novo J (2015) Analysis of the pressure stabilized Petrov-Galerkin method for the evolutionary Stokes equations avoiding time step restrictions. SIAM J Numer Anal 53:1005–1031
John V, Rang J (2010) Adaptive time step control for the incompressible Navier-Stokes equations. Comput Methods Appl Mech Eng 199:514–524
John V, Roland M (2007) Simulations of the turbulent channel flow at $\mathrm{Re}_\tau = 180$ with projection-based finite element variational multiscale methods. Int J Numer Methods Fluids 55:407–429
John V, Tobiska L (2000) Numerical performance of smoothers in coupled multigrid methods for the parallel solution of the incompressible Navier-Stokes equations. Int J Numer Methods Fluids 33:453–473
John V, Maubach JM, Tobiska L (1997) Nonconforming streamline-diffusion-finite-element-methods for convection-diffusion problems. Numer Math 78:165–188
John V, Tabata M, Tobiska L (1998) Error estimates for nonconforming finite element approximations of drag and lift in channel flows. Preprint 3/98. Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg
John V, Knobloch P, Matthies G, Tobiska L (2002) Non-nested multi-level solvers for finite element discretisations of mixed problems. Computing 68:313–341
John V, Matthies G, Rang J (2006a) A comparison of time-discretization/linearization approaches for the incompressible Navier-Stokes equations. Comput Methods Appl Mech Eng 195:5995–6010
John V, Kaya S, Layton W (2006b) A two-level variational multiscale method for convection-dominated convection-diffusion equations. Comput Methods Appl Mech Eng 195:4594–4603
John V, Layton W, Manica CC (2007) Convergence of time-averaged statistics of finite element approximations of the Navier-Stokes equations. SIAM J Numer Anal 46:151–179
John V, Kaya S, Kindl A (2008) Finite element error analysis for a projection-based variational multiscale method with nonlinear eddy viscosity. J Math Anal Appl 344:627–641
John V, Kindl A, Suciu C (2010) Finite element LES and VMS methods on tetrahedral meshes. J Comput Appl Math 233:3095–3102
John V, Linke A, Merdon C, Neilan M, Rebholz LG (2016) On the divergence constraint in mixed finite element methods for incompressible flows. SIAM Review (in press)
Kamke E (1944) Differentialgleichungen. Lösungsmethoden und Lösungen. Band I. Gewöhnliche Differentialgleichungen. Mathematik und ihre Anwendungen in Physik und Technik. Band 18_1, 3d ed. Akademische Verlagsgesellschaft, Leipzig, pp xxvi+666
Kay D, Loghin D, Wathen A (2002) A preconditioner for the steady-state Navier-Stokes equations. SIAM J Sci Comput 24:237–256
Kay DA, Gresho PM, Griffiths DF, Silvester DJ (2010) Adaptive time-stepping for incompressible flow. II. Navier-Stokes equations. SIAM J Sci Comput 32:111–128
Kellogg RB, Osborn JE (1976) A regularity result for the Stokes problem in a convex polygon. J Funct Anal 21:397–431
Klouček P, Rys FS (1994) Stability of the fractional step Θ-scheme for the nonstationary Navier-Stokes equations. SIAM J Numer Anal 31:1312–1335
Knobloch P (2001) Uniform validity of discrete Friedrichs' inequality for general nonconforming finite element spaces. Numer Funct Anal Optim 22:107–126
Kolmogoroff A (1941) The local structure of turbulence in incompressible viscous fluid for very large Reynold's numbers. C R (Doklady) Acad Sci URSS (N.S.) 30:301–305
Kolmogorov AN, Fomīn SV (1975) Reelle Funktionen und Funktionalanalysis. VEB Deutscher Verlag der Wissenschaften, Berlin, p 534. Übersetzt von der dritten russischen Auflage von D. Freitag, H. Palme und B. Stöckert, Hochschulbücher für Mathematik, Band 78

Konshin IN, Olshanskii MA, Vassilevski YV (2015) ILU preconditioners for nonsymmetric saddle-point matrices with application to the incompressible Navier-Stokes equations. SIAM J Sci Comput 37:A2171–A2197

Kraichnan RH (1967) Inertial ranges in two-dimensional turbulence. Phys Fluids 10:1417–1423

Král J, Wendland W (1986) Some examples concerning applicability of the Fredholm-Radon method in potential theory. Apl. Mat. 31:293–308

Kreiss H-O, Lorenz J (2004) Initial-boundary value problems and the Navier-Stokes equations. Classics in applied mathematics, vol 47. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, pp xviii+402. Reprint of the 1989 edition

Ladyženskaja OA (1967) New equations for the description of the motions of viscous incompressible fluids, and global solvability for their boundary value problems. Trudy Mat Inst Steklov 102:85–104

Ladyzhenskaya OA (1969) The mathematical theory of viscous incompressible flow. Second English edition, revised and enlarged. Translated from the Russian by Richard A. Silverman and John Chu. Mathematics and its Applications, vol 2. Gordon and Breach, Science Publishers, New York/London/Paris, pp xviii+224

Layton W (2002) A connection between subgrid scale eddy viscosity and mixed methods. Appl Math Comput 133:147–157

Layton W (2008) Introduction to the numerical analysis of incompressible viscous flows. Computational science & engineering, vol 6. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, pp xx+213. With a foreword by Max Gunzburger

Layton W, Lewandowski R (2003) A simple and stable scale-similarity model for large eddy simulation: energy balance and existence of weak solutions. Appl Math Lett 16:1205–1209

Layton W, Lewandowski R (2006) On a well-posed turbulence model. Discrete Contin Dyn Syst Ser B 6:111–128 (electronic)

Layton WJ, Rebholz LG (2012) Approximate deconvolution models of turbulence. Analysis, phenomenology and numerical analysis. Lecture notes in mathematics, vol 2042. Springer, Heidelberg, pp viii+184

Layton W, Tobiska L (1998) A two-level method with backtracking for the Navier-Stokes equations. SIAM J Numer Anal 35:2035–2054 (electronic)

Layton W, Manica CC, Neda M., Rebholz LG (2008) Numerical analysis and computational testing of a high accuracy Leray-deconvolution model of turbulence. Numer. Methods Partial Differ Equ 24:555–582

Layton W, Manica CC, Neda M, Rebholz LG (2010) Numerical analysis and computational comparisons of the NS-alpha and NS-omega regularizations. Comput Methods Appl Mech Eng 199:916–931

Lederer PL (2016) Pressure Robust Discretizations for Navier Stokes Equations: Divergence-free Reconstruction for Taylor-Hood Elements and High Order Hybrid Discontinuous Galerkin Methods. Diplomarbeit, Technische Universität Wien

Leonard A (1975) Energy cascade in large-eddy simulations of turbulent fluid flows. In: Frenkiel F, Munn R (eds) Turbulent diffusion in environmental pollution proceedings of a symposium held at Charlottesville. Advances in geophysics, vol 18, part A. Academic, New York, pp 237–248

Leray J (1933) Étude de diverses équations intégrales non linéaires et de quelques problèmes de l'hydrodynamique. J Math Pures Appl (9) 12:1–82

Leray J (1934a) Essai sur les mouvements plans d'une liquide visqueux que limitent des parois. J Math Pures Appl (9) 13:331–418

Leray J (1934b) Sur le mouvement d'un liquide visqueux emplissant l'espace. Acta Math. 63:193–248

Lesieur M (1997) Turbulence in fluids. Fluid mechanics and its applications, vol 40, 3rd edn. Kluwer Academic Publishers Group, Dordrecht, pp xxxii+515

Lesieur M, Metais O, Comte P (2005) Large-eddy simulations of turbulence. Cambridge University Press, New York, pp xii+219. With a preface by James J. Riley

Lilly D (1967) The representation of small-scale turbulence in numerical simulation experiments. In: Goldstine H (ed) Proceedings of the IBM scientific computing symposium on environmental sciences. IBM, Yorktown Heights, NY, pp 195–210

Lilly DK (1992) A proposed modification of the germano subgrid-scale closure method. Phys Fluids A 4:633–635

Linke A (2014) On the role of the Helmholtz decomposition in mixed methods for incompressible flows and a new variational crime. Comput Methods Appl Mech Eng 268:782–800

Linke A, Merdon C, Wollner W (2016a) Optimal L^2 velocity error estimate for a modified pressure-robust Crouzeix–Raviart Stokes element. IMA J Numer Anal (in press)

Linke A, Matthies G, Tobiska L (2016b) Robust arbitrary order mixed finite element methods for the incompressible Stokes equations with pressure independent velocity errors. ESAIM Math Model Numer Anal 50:289–309

Lions J-L, Magenes E (1972) Non-homogeneous boundary value problems and applications. Vol. I. Springer, New York/Heidelberg, pp xvi+357. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181

Lube G, Rapin G (2006) Residual-based stabilized higher-order FEM for a generalized Oseen problem. Math Models Methods Appl Sci 16:949–966

Lube G, Tobiska L (1990) A nonconforming finite element method of streamline diffusion type for the incompressible Navier-Stokes equations. J Comput Math 8:147–158

Lube G, Rapin G, Löwe J (2008) Local projection stabilization for incompressible flows: equal-order vs. inf-sup stable interpolation. Electron Trans Numer Anal 32:106–122

Manica CC, Merdan SK (2007) Finite element error analysis of a zeroth order approximate deconvolution model based on a mixed formulation. J Math Anal Appl 331:669–685

Mardal K-A, Schöberl J, Winther R (2013) A uniformly stable Fortin operator for the Taylor-Hood element. Numer Math 123:537–551

Marion M, Temam R (1998) Navier-Stokes equations: theory and approximation. In: Handbook of numerical analysis, vol VI. North-Holland, Amsterdam, pp 503–688

Marsden JE, Shkoller S (2001) Global well-posedness for the Lagrangian averaged Navier-Stokes (LANS-α) equations on bounded domains. R Soc Lond Philos Trans Ser A Math Phys Eng Sci 359:1449–1468. Topological methods in the physical sciences (London, 2000)

Marsden JE, Shkoller S (2003) The anisotropic Lagrangian averaged Euler and Navier-Stokes equations. Arch Ration Mech Anal 166:27–46

Matthies G (2001) Mapped finite elements on hexahedra. Necessary and sufficient conditions for optimal interpolation errors. Numer Algorithms 27:317–327

Matthies G, Tobiska L (2002) The inf-sup condition for the mapped Q_k-P_{k-1}^{disc} element in arbitrary space dimensions. Computing 69:119–139

Matthies G, Tobiska L (2005) Inf-sup stable non-conforming finite elements of arbitrary order on triangles. Numer Math 102:293–309

Matthies G, Tobiska L (2015) Local projection type stabilization applied to inf-sup stable discretizations of the Oseen problem. IMA J Numer Anal 35:239–269

Matthies G, Skrzypacz P, Tobiska L (2007) A unified convergence analysis for local projection stabilisations applied to the Oseen problem. M2AN Math. Model. Numer. Anal. 41:713–742

Matthies G, Lube G, Röhe L (2009) Some remarks on residual-based stabilisation of inf-sup stable discretisations of the generalised Oseen problem. Comput Methods Appl Math 9:368–390

Maxwell J (1879) On stresses in rarified gases arising from inequalities of temperature Philos Trans R Soc 170:231–256

Meneveau C, Katz J (2000) Scale-invariance and turbulence models for large-eddy simulation. Ann Rev Fluid Mech 32:1–32. Annual Reviews, Palo Alto, CA

Miles WW, Rebholz LG (2010) An enhanced-physics-based scheme for the NS-α turbulence model. Numer Methods Partial Differ Equ 26:1530–1555

Moser RD, Kim J, Mansour NN (1999) Direct numerical simulation of turbulent channel flow up to $Re_\tau = 590$. Phys Fluids 11:943–945

Müller-Urbaniak S (1993) Eine Analyse des Zwischenschritt-θ-Verfahrens zur Lösung der instationären Navier-Stokes-Gleichungen. PhD thesis, Univ. Heidelberg, Naturwiss.-Math. Gesamtfak., Heidelberg, pp. 78

Nabh G (1998) On higher order methods for the stationary incompressible Navier-Stokes equations. PhD thesis, Univ. Heidelberg, Naturwissenschaftlich-Mathematische Gesamtfakultät, Heidelberg, pp 104

Navier C (1823) Mémoire sur les lois du mouvement des fluiales. Mém Acad R Soc 6:389–440

Neilan M (2015) Discrete and conforming smooth de Rham complexes in three dimensions. Math Comp 84:2059–2081

Nirenberg L (1959) On elliptic partial differential equations. Ann Scuola Norm Sup Pisa (3) 13:115–162

Nitsche J (1968) Ein Kriterium für die Quasi-Optimalität des Ritzschen Verfahrens. Numer Math 11:346–348

Ohmori K, Ushijima T (1984) A technique of upstream type applied to a linear nonconforming finite element approximation of convective diffusion equations. RAIRO Anal Numér 18:309–332

Olshanskii MA (2002) A low order Galerkin finite element method for the Navier-Stokes equations of steady incompressible flow: a stabilization issue and iterative methods. Comput Methods Appl Mech Eng 191:5515–5536

Olshanskii MA, Reusken A (2004) Grad-div stabilization for Stokes equations. Math Comp 73:1699–1718

Olshanskii MA, Tyrtyshnikov EE (2014) Iterative methods for linear systems. Theory and applications. Society for Industrial and Applied Mathematics, Philadelphia, PA, pp xvi+247

Olshanskii MA, Vassilevski YV (2007) Pressure Schur complement preconditioners for the discrete Oseen problem. SIAM J Sci Comput 29:2686–2704

Olshanskii M, Lube G, Heister T, Löwe J (2009) Grad-div stabilization and subgrid pressure models for the incompressible Navier-Stokes equations. Comput Methods Appl Mech Eng 198:3975–3988

Parés C (1992) Existence, uniqueness and regularity of solution of the equations of a turbulence model for incompressible fluids. Appl Anal 43:245–296

Patankar SV (1980) Numerical heat transfer and fluid flow. Series in computational methods in mechanics and thermal sciences. Hemisphere Publishing Corporation, Washington, New York, London/McGraw-Hill Book Company, New York, pp XIII+197

Piomelli U (1999) Large-eddy simulation: achievements and challenges. Progr Aerosp Sci 35:335–362

Piomelli U, Balaras E (2002) Wall-layer models for large-eddy simulations. Ann Rev Fluid Mech 34:349–374. Annual Reviews, Palo Alto, CA

Pope SB (2000) Turbulent flows. Cambridge University Press, Cambridge, pp xxxiv+771

Principe J, Codina R, Henke F (2010) The dissipative structure of variational multiscale methods for incompressible flows. Comput Methods Appl Mech Eng 199:791–801

Prohl A (1997) Projection and quasi-compressibility methods for solving the incompressible Navier-Stokes equations. Advances in numerical mathematics. B. G. Teubner, Stuttgart, pp xiv+294

Qin J (1994) On the convergence of some low order mixed finite elements for incompressible fluids. PhD thesis, Department of Mathematics, Pennsylvania State University

Rang J (2008) Pressure corrected implicit θ-schemes for the incompressible Navier-Stokes equations. Appl Math Comput 201:747–761

Rannacher R (1992) On Chorin's projection method for the incompressible Navier-Stokes equations. In: The Navier-Stokes equations II—theory and numerical methods (Oberwolfach, 1991). Lecture notes in mathematics, vol 1530. Springer, Berlin, pp 167–183

Rannacher R, Turek S (1992) Simple nonconforming quadrilateral Stokes element. Numer Methods Partial Differ Equ 8:97–111

Rasthofer U (2015) Computational multiscale methods for turbulent single and two-phase flows. Bericht 27. PhD thesis, Lehrstuhl für Numerische Mechanik, Technische Universität München

Rasthofer U, Gravemeier V (2013) Multifractal subgrid-scale modeling within a variational multiscale method for large-eddy simulation of turbulent flow. J Comput Phys 234:79–107
Raviart P-A, Thomas JM (1977) A mixed finite element method for 2nd order elliptic problems. Mathematical aspects of finite element methods (proceedings of the conference, Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975). Lecture notes in mathematics, vol 606. Springer, Berlin, pp 292–315
Richardson LF (1922) Weather prediction by numerical process. Cambridge University Press, Cambridge
Rockel S (2013) Über Formen des konvektiven Terms in Finite–Elemente–Diskretisierungen der inkompressiblen Navier–Stokes–Gleichungen. Diplomarbeit, Freie Universität Berlin
Rodi W, Ferziger JH, Breuer M, Pourquié M (1997) Status of large eddy simulation: results of a workshop. ASME J Fluids Eng 119:248–262
Röhe L, Lube G (2010) Analysis of a variational multiscale method for large-eddy simulation and its application to homogeneous isotropic turbulence. Comput Methods Appl Mech Eng 199:2331–2342
Roos H-G, Stynes M, Tobiska L (2008) Robust numerical methods for singularly perturbed differential equations. Convection-diffusion-reaction and flow problems. Springer series in computational mathematics, vol 24, 2nd edn. Springer, Berlin, pp xiv+604
Rudin W (1991) Functional analysis. International series in pure and applied mathematics, 2nd edn. McGraw-Hill, Inc., New York, pp xviii+424
Saad Y (1993) A flexible inner-outer preconditioned GMRES algorithm. SIAM J Sci Comput 14:461–469
Saad Y (2003) Iterative methods for sparse linear systems, 2nd edn. Society for Industrial and Applied Mathematics, Philadelphia, PA, pp xviii+528
Saad Y, Schultz MH (1986) GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM J Sci Stat Comput 7:856–869
Sagaut P (2006) Large eddy simulation for incompressible flows. In: Scientific computation, 3rd edn. Springer, Berlin, pp xxx+556. An introduction, translated from the 1998 French original, with forewords by Marcel Lesieur and Massimo Germano, with a foreword by Charles Meneveau
Schäfer M, Turek S (1996) Benchmark computations of laminar flow around a cylinder (with support by F. Durst, E. Krause and R. Rannacher). In: Flow simulation with high-performance computers II. DFG priority research programme results 1993–1995. Vieweg, Wiesbaden, pp 547–566
Schenk O, Bollhöfer M, Römer RA (2008) On large-scale diagonalization techniques for the anderson model of localization SIAM Rev 50:91–112
Schieweck F (1997) Parallele Lösung der stationären inkompressiblen Navier-Stokes Gleichungen. Habilitation thesis, Magdeburg: Univ. Magdeburg, Fakultät Mathematik, pp 142
Schieweck F (2000) A general transfer operator for arbitrary finite element spaces. Preprint 00-25. Fakultät für Mathematik, Otto-von-Guericke-Universität Magdeburg
Schieweck F, Tobiska L (1989) A nonconforming finite element method of upstream type applied to the stationary Navier-Stokes equation. RAIRO Modél Math Anal Numér 23:627–647
Schieweck F, Tobiska L (1996) An optimal order error estimate for an upwind discretization of the Navier-Stokes equations. Numer Methods Partial Differ Equ 12:407–421
Schönknecht N (2015) On solvers for saddle point problems arising in finite element discretizations of incompressible flow problems. Master thesis, Freie Universität Berlin
Scott LR, Vogelius M (1985) Conforming finite element methods for incompressible and nearly incompressible continua. Large-scale computations in fluid mechanics, Part 2 (La Jolla, CA, 1983). Lectures in applied mathematics, vol 22. American Mathematical Society, Providence, RI, pp 221–244
Scott LR, Zhang S (1990) Finite element interpolation of nonsmooth functions satisfying boundary conditions. Math Comp 54:483–493

Serrin J (1963) The initial value problem for the Navier-Stokes equations. In: Nonlinear problems (proceedings of the sympos., Madison, WI, 1962). University of Wisconsin Press, Madison, WI, pp 69–98

Shakib F, Hughes TJR, Johan Z (1991) A new finite element formulation for computational fluid dynamics. X. The compressible Euler and Navier-Stokes equations. Comput Methods Appl Mech Eng 89:141–219. Second world congress on computational mechanics, part I (Stuttgart, 1990)

Shen J (1992a) On error estimates of projection methods for Navier-Stokes equations: first-order schemes. SIAM J Numer Anal 29:57–77

Shen J (1992b) On error estimates of some higher order projection and penalty-projection methods for Navier-Stokes equations. Numer Math 62:49–73

Shen J (1996) On error estimates of the projection methods for the Navier-Stokes equations: second-order schemes. Math Comp 65:1039–1065

Smagorinsky J (1963) General circulation experiments with the primitive equations. Mon Weather Rev 91:99–164

Sohr H (1983) Zur Regularitätstheorie der instationären Gleichungen von Navier-Stokes. Math Z 184:359–375

Sohr H (2001) The Navier-Stokes equations. An elementary functional analytic approach. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser Verlag, Basel, pp x+367

Sonneveld P (1989) CGS, a fast Lanczos-type solver for nonsymmetric linear systems. SIAM J Sci Stat Comput 10:36–52

Stanculescu I (2008) Existence theory of abstract approximate deconvolution models of turbulence. Ann Univ Ferrara Sez VII Sci Mat 54:145–168

Stenberg R (1984) Analysis of mixed finite elements methods for the Stokes problem: a unified approach. Math Comp 42:9–23

Stenberg R (1987) On some three-dimensional finite elements for incompressible media. Comput Methods Appl Mech Eng 63:261–269

Stenberg R (1990) Error analysis of some finite element methods for the Stokes problem. Math Comp 54:495–508

Stolz S, Adams NA (1999) An approximate deconvolution procedure for large-eddy simulation. Phys Fluids 11:1699–1701

Stolz S, Adams NA, Kleiser L (2001) An approximate deconvolution model for large-eddy simulation with application to incompressible wall-bounded flows. Phys Fluids 13:997–1015

Stroud AH (1971) Approximate calculation of multiple integrals. Prentice-Hall series in automatic computation. Prentice-Hall, Inc., Englewood Cliffs, NJ, pp xiii+431

Stüben K (2001) A review of algebraic multigrid. J Comput Appl Math 128:281–309. Numerical analysis 2000, vol VII. Partial differential equations

Świerczewska A (2006) A dynamical approach to large eddy simulation of turbulent flows: existence of weak solutions. Math Methods Appl Sci 29:99–121

Tabata M, Tagami D (2000) Error estimates for finite element approximations of drag and lift in nonstationary Navier-Stokes flows. Jpn J Ind Appl Math 17:371–389

Taylor G (1923) On the decay of vortices in viscous fluid. Philos Mag XLVI:671–674

Témam R (1969) Sur l'approximation de la solution des équations de Navier-Stokes par la méthode des pas fractionnaires. II. Arch Ration Mech Anal 33:377–385

Temam R (1982) Behaviour at time $t = 0$ of the solutions of semilinear evolution equations. J Differ Equ 43:73–92

Temam R (1984) Navier-Stokes equations. Theory and numerical analysis. Studies in mathematics and its applications, vol 2, 3rd edn. North-Holland Publishing Co., Amsterdam, pp xii+526. With an appendix by F. Thomasset

Temam R (1986) Infinite-dimensional dynamical systems in fluid mechanics. In: Nonlinear functional analysis and its applications, part 2 (Berkeley, CA, 1983). Proceedings of the symposium. Pure Mathematics, vol 45. American Mathematical Society, Providence, RI, pp 431–445

Temam R (1995) Navier-Stokes equations and nonlinear functional analysis. CBMS-NSF regional conference series in applied mathematics, vol 66, 2nd edn. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, pp xiv+141

Temam R (1997) Infinite-dimensional dynamical systems in mechanics and physics. Applied mathematical sciences, vol 68, 2nd edn. Springer, New York, pp xxii+648

Tiesinga G, Wubs FW, Veldman AEP (2002) Bifurcation analysis of incompressible flow in a driven cavity by the Newton-Picard method. J Comput Appl Math 140:751–772

Timmermans L, Minev P, van de Vosse F (1996) An approximate projection scheme for incompressible flow using spectral elements. Int J Numer Methods Fluids 22:673–688

Tobiska L, Lube G (1991) A modified streamline diffusion method for solving the stationary Navier-Stokes equation. Numer Math 59:13–29

Tobiska L, Verfürth R (1996) Analysis of a streamline diffusion finite element method for the Stokes and Navier-Stokes equations. SIAM J Numer Anal 33:107–127

Triebel H (1972) Höhere analysis. VEB Deutscher Verlag der Wissenschaften, Berlin, p 704. Hochschulbücher für Mathematik, Band 76

Trilinos (2016) https://trilinos.org. Accessed 22 June 2016

Trottenberg U, Oosterlee CW, Schüller A (2001) Multigrid. Academic Press, Inc., San Diego, CA, pp xvi+631. With contributions by A. Brandt, P. Oswald and K. Stüben

Turek S (1999) Efficient solvers for incompressible flow problems. An algorithmic and computational approach. Lecture notes in computational science and engineering, vol 6. Springer, Berlin, pp xvi+352. With 1 CD-ROM (“Virtual Album”: UNIX/LINUX, Windows, POWERMAC; “FEATFLOW 1.1”: UNIX/LINUX)

Umla R (2009) Stabilisierte Finite–Element Verfahren für die Konvektions-Diffusions-Gleichungen und Oseen-Gleichungen. Diplomarbeit, Universität des Saarlandes, FR 6.1 – Mathematik

ur Rehman M, Vuik C, Segal G (2008) A comparison of preconditioners for incompressible Navier-Stokes solvers. Int J Numer Methods Fluids 57:1731–1751

van Cittert P (1931) Zum Einfluß der Spaltbreite auf die Intensitätsverteilung in Spektrallinien. II. Zeitschrift für Physik 69:298–308

van der Bos F, Geurts BJ (2005) Commutator errors in the filtering approach to large-eddy simulation. Phys Fluids 17:035108, 20

van der Vorst HA (1992) Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems. SIAM J Sci Statist Comput 13:631–644

van Driest E (1956) On turbulent flow near a wall. J Aeronaut Sci 23:1007–1011, 1036

van Kan J (1986) A second-order accurate pressure-correction scheme for viscous incompressible flow. SIAM J Sci Stat Comput 7:870–891

Vaněk P, Mandel J, Brezina M (1996) Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems. Computing 56:179–196. International GAMM-workshop on multi-level methods (Meisdorf, 1994)

Vanka SP (1986) Block-implicit multigrid solution of Navier-Stokes equations in primitive variables. J Comput Phys 65:138–158

Vassilevski PS, Lazarov RD (1996) Preconditioning mixed finite element saddle-point elliptic problems. Numer Linear Algebra Appl 3:1–20

Verchota GC, Vogel AL (2003) A multidirectional Dirichlet problem. J Geom Anal 13:495–520

Verfürth R (1984) Error estimates for a mixed finite element approximation of the Stokes equations. RAIRO Anal Numér 18:175–182

Verfürth R (1989) A posteriori error estimators for the Stokes equations. Numer Math 55:309–325

Verfürth R (1994) A posteriori error estimation and adaptive mesh-refinement techniques. J Comput Appl Math 50:67–83

Verfürth R (1996) A review of a posteriori error estimation and adaptive mesh-refinement techniques. Wiley, Chichester; B. G. Teubner, Stuttgart, pp vi+127

Verfürth R (2013) A posteriori error estimation techniques for finite element methods. Numerical mathematics and scientific computation. Oxford University Press, Oxford, pp xx+393

Winckelmans G, Jeanmart H, Carati D (2002) On the comparison of turbulence intensities from large-eddy simulation with those from experiment or direct numerical simulation. Phys Fluids 14:3

Xu J, Zikatanov L (2003) Some observations on Babuška and Brezzi theories. Numer Math 94:195–202

Yosida K (1995) Functional analysis. Classics in mathematics. Springer, Berlin, pp xii+501. Reprint of the sixth (1980) edition

Zhang S (2005) A new family of stable mixed finite elements for the 3D Stokes equations. Math Comp 74:543–554

Zhang S (2009) Bases for C0-P1 divergence-free elements and for C1-P2 finite elements on union jack grids. http://www.math.udel.edu/~szhang/research/p/uj.pdf. Accessed 20 July 2016

Zang Y, Street RL, Koseff JR (1993) A dynamic mixed subgrid-scale model and its application to turbulent recirculating flows. Phys Fluids A 5:3186–3196

Index of Subjects

V. John, *Finite Element Methods for Incompressible Flow Problems*, Springer Series in Computational Mathematics 51, DOI 10.1007/978-3-319-45750-5