Numerical Methods of Continuum Mechanics

Version as of August 24, 2015

Lecture notes
by
Prof. Dr. Vincent Heuveline
Faculty of Mathematics and Computer Science
Heidelberg University

Summer term 2015

Contents

1 Introduction **5**
1.1 Examples for fields of applications . . . 5
1.1.1 Elasticity and plasticity . . . 5
1.1.2 Fluid mechanics . . . 6
1.1.3 Combination of elasticity/plasticiy and fluid mechanics . . . 9
1.2 Notation . . . 10
1.3 Review of Vector Calculus . . . 10
1.4 Path of lecture . . . 12

2 The equation of fluid flow - Introduction to modelling **13**
2.1 Lagrangian and Eulerian Systems: Substantial Derivative . . . 13
2.1.1 The Lagrangian viewpoint . . . 13
2.1.2 The Eulerian viewpoint . . . 14
2.1.3 The substantial derivative . . . 14
2.2 Conservation of mass . . . 15
2.3 Conservation of momentum . . . 16
2.3.1 Conservation of angular momentum . . . 17
2.3.2 Conservation of energy . . . 18
2.3.3 Balance equations . . . 20
2.3.4 Kinematic properties . . . 20
2.3.5 Material properties . . . 27

3 Approximating Steady Flows **33**
3.1 The Stokes problem: Introduction to Mixed Methods . . . 33
3.2 Formulation and Stability of the Approximation . . . 37

4 Time dependent Navier-Stokes equations **43**

5 Saddle-point problem **47**
5.1 General setting . . . 47
5.1.1 Discrete setting . . . 48
5.2 The inf – sup condition . . . 49
5.2.1 The Stokes equations . . . 49
5.2.2 Laplace equation . . . 49
5.2.3 The inf – sup condition . . . 50

6 A posteriori error estimation **55**
6.1 Introduction . . . 55
6.2 Residual based error estimator . . . 55
6.2.1 Residual estimators . . . 57
6.2.2 Dual weighted Residual method (DWR) . . . 58
6.3 Exact grid optimization . . . 60

7 Practical aspects **61**
7.1 Solution of the discrete Stokes problems 61
7.1.1 Schur complement method 62
7.2 Solution of the stationary Navier-Stokes equations 64
7.2.1 Discretization of the convective terms 64
7.2.2 Linearization 71
7.2.3 Algebraic solution of the linearized problems 72
7.3 Solution of the instationary Navier-Stokes equations 75
7.3.1 Time-stepping schemes 75
7.3.2 Projection methods 77

Bibliography **79**

Index **81**

1 Introduction

Continuum mechanics deals with mechanics of continuous media. The emphasis lies on *continuous*, which means that the considered media are *not* understood as a set of point masses as their atomic structure suggests to, but the medium is seen as a *continuum*. Consequently, observations are described from a macroscopic point of view and not a microscopic one. This offers the possibility to mathematically describe the quantities of interest, for example the velocity of the medium or its deformation, as - often continuous or even continuously differentiable - functions and analyse them by means of calculus. This is known as *continuum hypothesis*.

By *modelling* the underlying physical principles this approach naturally leads to *partial differential equations*, whose analytical solution for practically relevant applications cannot be given in an explicit way. As a consequence, to examine these processes, it is necessary to develop adequat numerical methods and to analyse both the methods themself and their results.

First, a short, for sure not complete, review will be given concerning the field of applications of continuum mechanics. Then some basic and regularly used notations will be introduced.

1.1 Examples for fields of applications

1.1.1 Elasticity and plasticity

In elasticity and plasticity one deals with elastic and plastic deformations of a medium under the influence of forces. An elastically deformed medium returns to its original state after removement of the forces, whereas a medium that is plastically deformed will *not* return to ist primary state, but stay deformed. Whether a deformation is elastic or plastic depends on many factors, for example the material properties, the strength of the operating forces and additionally ongoing processes.

Example 1.1 As an example to illustrate the difference between elastic and plastic deformation, we choose a coil spring. After having expanded the spring by an external force, two scenarios can arise after the removement of the force:

- If the springs deformation was „sufficiently small“, the spring will contract again and return to its original state. So the spring has been *elastically* deformed in this situation.
- If the spring is strechted very strong, so that the metal forms just a „long thread“, it will - depending on the material - not contract completely and stay deformed. So the deformation is a *plastic* one.

To demonstrate the relevance of additional, parallel running processes, we assume the spring to be made of a so called shape-memory material. Now, in the second scenario, too, the spring will return to its primary (deformation) state, when it is sufficiently heated after the force has been removed. Considering deformation and heating as *one* net process, the shape-memory spring has been *elastically* deformed.

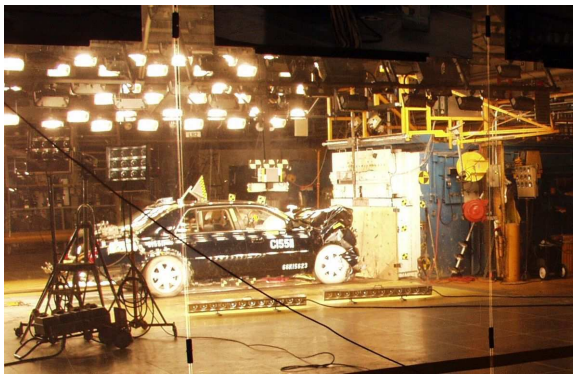

Figure 1.1: Crash test [1].

Of course, this is only a simple example, nonetheless the fundamental phenomena play an important role in applications:

- *Structural analysis*: Can a building bear mechanical stress, for example due to the weight of the furniture or a storm? Or will plastic deformations occur, like ruptures in the ceiling/floor or outer walls?
- *Medicine*: simulation of cutting tissues during a surgery. This is a current object of research in our group.
- *Crash test*, for example in automobile industry.

1.1.2 Fluid mechanics

Fluid mechanics deals with the motion of *fluids*, for example water, air or honey, and the forces acting in this context. In particular, not only external forces like gravity, but also internal forces are important, as for example friction.

Possible fields of applications can of course be found in every situation in which one is interested in the fluids motion:

- *Environmental sciences*: weather forecast and climate prediction (atmosphere and ocean, i.e. a coupled system of air *and* water), pollutant dispersion (eg. infiltration in the ground). Particularly, also *assurance companies* are interested in the results of such calculations.
- *Medicine*: simulation of the respiratory system or blood flow
- *Traffic planning*: modelling of the traffic as a flowing fluid
- *Space flight*: construction and simulation of rocket engines
- *Aviation*: air circulation around a plane. In particular the flow profile near the wing is important for sufficient bouyancy and a secure flight, but also the wake turbulences forming behind the plane. These negatively effect the security of the following plane seriously!
- *Automobile industry*: air circulation around a car. The resistance of the autobody to the air flow is relevant for both the driving safety (force pressing the car to the ground) and the fuel consumption.

Figure 1.2: Currents in the environment [2].

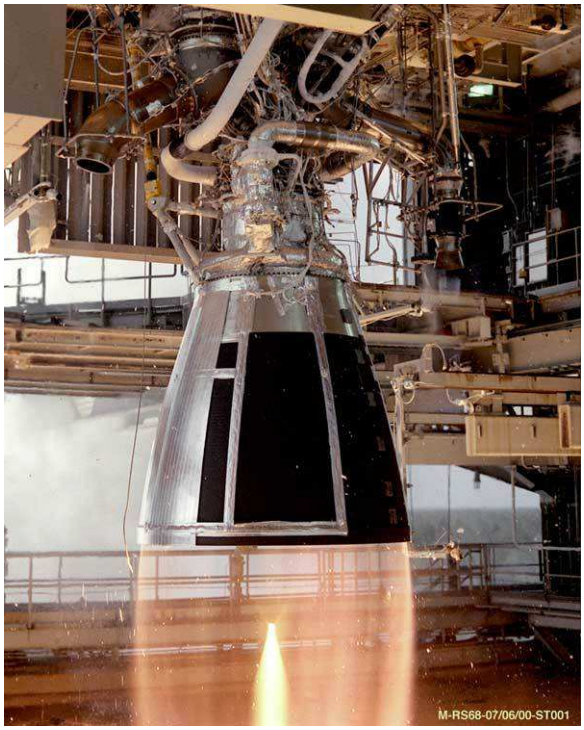

Figure 1.3: Rocket engine [3].

Figure 1.4: Wake turbulence [4].

Figure 1.5: Measuring the aeroacoustics in a wind tunnel [5].

Furthermore, during the process of constructing aircrafts and cars appear interesting couplings with optimisation theory. One can for instance try to calculate the design of the autobody or wing in such a way, that certain conditions, eg. a force strong enough to keep the car on the ground or to raise the plane into the air, are fulfilled and at the same time the aerodynamic drag (and hence the fuel consumption) is minimised.

1.1.3 Combination of elasticity/plasticiy and fluid mechanics

Also due to the massively increased computing power in the last 20 years, applications which combine elastic/plastic and fluid mechanical phenomena are getting more and more important. In the above examples they were separated, meaning for instance the geometry of a body in a fluid flow was considered as given for the simulation of the flow. However, there are scenarios, that require a coupled consideration of these phenomena, and in this case you are talking about *fluid-structure interaction (FSI)*.

Example 1.2 For the planning of a wind farm the wind turbines must be constructed and positioned in such a way that they will not be damaged at work - in the worst case they fall over or the rotor blades break. It is important to remember that the turbine not only need to bear the maximum wind force usually occuring for the region, but it heavily influences the air flow in the wind farm itself. Behind the rotor blades, similar to the aircraft wings, severe air turbulences arise. By placing two wind turbines too close or in a disadvantageous orientation with respect to each other, these vortices can carry enough energy with them to damage an other wind turbine! For this reason, in such a scenario the wind flow *and* the elastic/plastic deformation of the wind turbines should be simulated and calculated *together*, i.e. completely *coupled*. Again, a combination with optimisation methods is possible, concerning the construction of the wind turbines or their placement in the wind farm.

Figure 1.6: Wind farm [6].

1.2 Notation

We consider open bounded domains $\Omega \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, with boundary $\partial\Omega$. In general, the unknown function is denoted by $u(t)$, $u(x)$ or $u(x, t)$ for function arguments $x \in \mathbb{R}^d$ and $t \in (0, T)$, respectively. The variable t typically denotes a time variable and $x = (x_1, \ldots, x_d)^\top \in \mathbb{R}^d$ a spatial variable. For functions $u(t)$, $u(x)$ or $u(x, t)$ we denote total derivatives and partial derivatives, respectively, by

$$\mathrm{d}_t u := \frac{\mathrm{d}u}{\mathrm{d}t}, \quad \partial_t u := \frac{\partial u}{\partial t}, \quad \partial_i u := \frac{\partial u}{\partial x_i}.$$

The *gradient of a scalar function* and the *divergence of a vector function* are defined as:

$$\begin{aligned} \operatorname{grad} u := \quad & \nabla u := \quad (\partial_1 u, \ldots, \partial_d u)^\top, \\ \operatorname{div} u := \quad & \nabla \cdot u := \quad \partial_1 u_1 + \cdots + \partial_d u_d. \end{aligned}$$

In three dimensions, we define the *curl (rotation) of a vector function* as

$$\operatorname{rot} u := \nabla \times u := \begin{pmatrix} \partial_2 u_3 - \partial_3 u_2 \\ \partial_3 u_1 - \partial_1 u_3 \\ \partial_1 u_2 - \partial_2 u_1 \end{pmatrix}$$

The operator $\nabla = (\partial_1, \ldots, \partial_d)^\top$ is the so called *nabla operator*. Combining divergence and gradient operator gives the *Laplace operator*

$$\Delta u := \operatorname{div}(\nabla u) = \partial_1^2 u + \cdots + \partial_d^2 u.$$

Furthermore, the directional derivative in direction of $n \in \mathbb{R}^d$ is denoted by $\partial_n u := n \cdot \nabla u$.

1.3 Review of Vector Calculus

Theorem 1.3 (Gauss or divergence theorem) *For any smooth vector field F over a region $\Omega \subset \mathbb{R}^3$ with a smooth boundary $S = \partial\Omega$, it holds*

$$\int_\Omega \operatorname{div} F \mathrm{d}V = \int_S F \cdot \mathbf{n} \mathrm{d}A.$$

PICTURE!!!

Figure 1.7: Gauss theorem

Theorem 1.4 (General transport theorem) *Let F be a smooth vector (or scalar) field on a region $R(t)$, whose boundary is $S(t)$, and let $\mathbf{U}$ be the velocity field of the time dependent movement of $S(t)$. Then,*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{R(t)} F(x,t)\mathrm{d}V = \int_{R(t)} \partial_t F \mathrm{d}V + \int_{S(t)} F\mathbf{U}\cdot\mathbf{n}\mathrm{d}S.$$

Proof $\Phi(x,t) = \Phi(x(\xi,t),t)$. Hereby, we consider the mapping

$$\xi \to x(\xi,t)$$

describing the trajectory of ξ.

$$\Delta(\xi,t) = \det\left(\frac{\partial x_i}{\partial \xi_j}\right)_{i,j=1,2,3} > 0.$$

Considering a reference volume $V_0 = V(0)$,

$$\int_{V(t)} \Phi(x,t)\mathrm{d}V = \int_{V_0} \Phi(x(\xi,t),t)\Delta(\xi,t)\mathrm{d}\xi.$$

Hence

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{V(t)} \Phi(x,t)\mathrm{d}V = \int_{V_0} \underbrace{\left\{\mathrm{d}_t\Phi(x(\xi,t),t)\Delta(\xi,t) + \Phi(x(\xi,t),t)\partial_t\Delta(\xi,t)\right\}}_{=\frac{\mathrm{d}}{\mathrm{d}t}(\Phi(x(\xi,t),t)\Delta(\xi,t))}\mathrm{d}\xi.$$

First term on right-hand side:

$$\mathrm{d}_t\Phi(x(\xi,t),t) = \partial_t\Phi(x,t) + \mathbf{U}(x,t)\cdot\nabla_x\Phi(x,t).$$

Second term on right-hand side: Assuming $a_{ij} = \partial_{\xi_j} x_i$ we obtain

$$\partial_t\Delta(\xi,t) = \partial_{a_{ij}}\Delta(\xi,t)\partial_t a_{ij} = \partial_{a_{ij}}\Delta\partial_{\xi_j}\mathbf{U}_i = \partial_{a_{ij}}\Delta\partial_k\mathbf{U}_i\partial_{\xi_j}x_k = \partial_{a_{ij}}\Delta\partial_k\mathbf{U}_i a_{kj}.$$

Therefore,

$$\partial_t\Delta(\xi,t) = \Delta_{ij}a_{kj}\partial_k\mathbf{U}_i = \delta_{ik}\Delta\partial_k\mathbf{U}_i = \Delta\partial_i\mathbf{U}_i = \Delta\mathrm{div}\ \mathbf{U}.$$

We obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{V(t)} \Phi(x,t)\mathrm{d}x = \int_{V_0}\left\{\partial_t\Phi + \mathbf{U}\cdot\nabla_x\Phi + \Phi\mathrm{div}\ \mathbf{U}\right\}\Delta(\xi,t)\mathrm{d}\xi = \int_{V(t)}\left\{\partial_t\Phi + \mathrm{div}\ \Phi\mathbf{U}\right\}\mathrm{d}x.$$

□

Theorem 1.5 (Reynolds transport theorem) *Let Φ be any smooth vector (or scalar) field, and suppose $R(t)$ is a fluid element with surface $S(t)$ travelling at the flow velocity $\mathbf{U}$. Then*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{R(t)} \Phi(x,t)\mathrm{d}V = \int_{R(t)} \partial_t\Phi(x,t)\mathrm{d}V + \int_{S(t)} \Phi(x,t)\mathbf{U}\cdot\mathbf{n}\mathrm{d}S.$$

Proof We prove the Reynolds transport theorem for the scalar case.

We begin by defining

$$F(x(t), y(t), z(t), t) := \int_{R(t)} \Phi \mathrm{d}V$$

with $R(t)$ taken to be a time-dependent fluid element (and hence the time dependence of the coordinate arguments of F). Now Theorem 1.4 applied to the scalar field Φ yields the result. □

1.4 Path of lecture

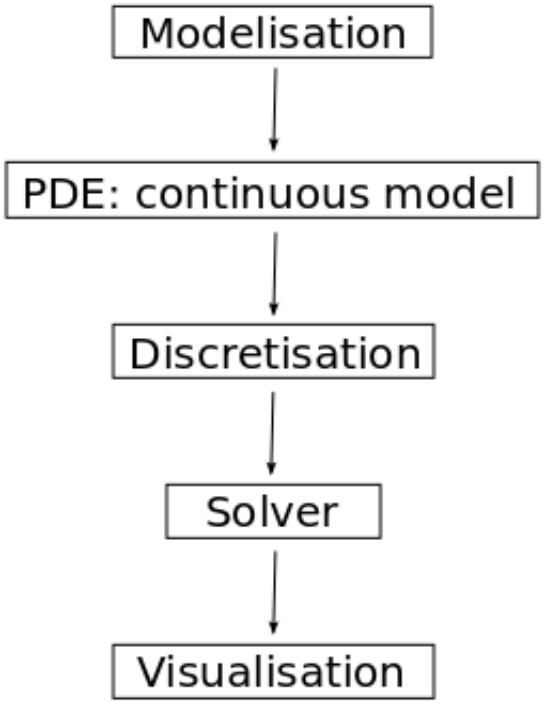

Figure 1.8: Simulation workflow

2 The equation of fluid flow - Introduction to modelling

In this chapter, we will especially adress aspects of modelling phenomena in fluid mechanics. In this context, it is essential to deal with the question, how to model mechanical properties of a fluid, so, as a consequence, also aspects of elasticity/plasticity will occur.

It shall be emphasised that this chapter is only an introduction to modelling; due to the complexity of the subject, many interesting and important aspects still are object of research and a detailed presentation is beyond the scope of the lecture.

We describe the state of a fluid by its state variables. To characterise the fluid flow, in general the following state variables need to be considered:

- velocity of the fluid $v(x,t)$,
- density of the fluid $\rho(x,t)$,
- temperature of the fluid $T(x,t)$ and
- pressure $p(x,t)$.

In the case of a three-dimensional flow, we need to model six equations to describe the six state variables. Depending on the physics of the considered problem, it can be well-motivated to assume some of the variables, for example density or temperature, to be constant or negligible.

The modelling is done by considering conservation principles.

2.1 Lagrangian and Eulerian Systems: Substantial Derivative

In the study of fluid motion there are two main approaches to describing the fluid behaviour: *Lagrangian* versus *Eulerian* viewpoint.

2.1.1 The Lagrangian viewpoint

The Lagrangian viewpoint consists of *following the material particles* of the continuum in their motion.

$$\begin{aligned} \varphi : \mathbb{R}^d \times [T_0, T_{final}] &\to \mathbb{R}^d \times [T_0, T_{final}] \\ (X,t) &\mapsto \varphi(X,t) = (x,t) \end{aligned}$$

The mapping φ allows to link X and x during time by the law of motion, namely

$$x = x(X,t), \quad t = t,$$

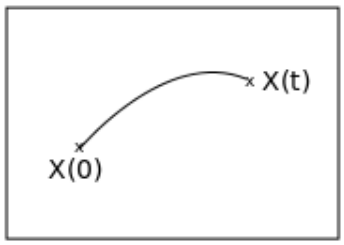

Figure 2.1: Lagrangian viewpoint

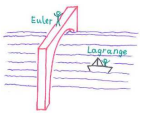

Figure 2.2: Lagrangian versus Eulerian viewpoint

which explicitly states the particular nature of φ: first, the spatial coordinate x depends on both the material particle X and time t, and, second, physical time is measured by the same variable t in both material and spatial domains.

Obviously, the one-to-one mapping φ must verify

$$\det\left(\frac{\partial x}{\partial X}\right) > 0$$

in order to impose a one-to-one correspondence (non zero) and to avoid change of orientation in the reference axes (positive). In order to obtain a complete description of the flow field, it is necessary to track a very large number of fluid particles. In engineering applications there is typically a need to know the fluid properties at a given point or location independently from the origin of the particles.

2.1.2 The Eulerian viewpoint

The Eulerian description corresponds to a coordinate system *fixed in space*, and in which fluid properties are studied as functions of time as the flow passes fixed spatial locations.

2.1.3 The substantial derivative

Definition 2.1 (Substantial derivative) *The* substantial derivative *of any fluid property*

$$f(x, y, z, t)$$

in a flow field with velocity $\mathbf{U} = (u, v, w)^\top$ *is given by*

$$\mathrm{d}_t f = \partial_t f + u\partial_x f + v\partial_y f + w\partial_z f = \partial_t f + \mathbf{U} \cdot \nabla f.$$

We recall, that the operator ∇ *is a vector differential operator* $\nabla = (\partial_x, \partial_y, \partial_z)^\top$.

(x, y, z) at time t in the Lagrangian viewpoint is defined as

$$\begin{aligned} x &= x(X, t), \\ y &= y(X, t), \\ z &= z(X, t). \end{aligned}$$

Therefore,

$$f\left(x\left(X,t\right), y\left(X,t\right), z\left(X,t\right), t\right)$$

and it follows

$$\begin{aligned} \mathrm{d}_t f &= \partial_x f \partial_t x + \partial_y f \partial_t y + \partial_z f \partial_t z + \partial_t f \\ &= \partial_x f u + \partial_y f v + \partial_z f w + \partial_t f \\ &= \partial_t f + \mathbf{U} \cdot \nabla f. \end{aligned}$$

The last term on the right hand side is related to the *transport* of the property f with the velocity field $\mathbf{U}$.

2.2 Conservation of mass

The equation describing the conservation of mass is called *continuity equation*. If mass is conserved, then the value of change of mass within a control volume $V \subset \mathbb{R}^d$ needs to be equal to the mass flux over the boundary ∂V (cf. Theorem 1.4):

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_V \rho \mathrm{d}x = - \int_{\partial V} (\rho \mathbf{v}) \cdot \mathbf{n} \mathrm{d}o$$

with the velocity vector $\mathbf{v}$ and the outward unit vector $\mathbf{n}$ perpendicular to ∂V.

The divergence theorem (Gauss's theorem) implies

$$\int_V \partial_t \rho + \operatorname{div}\ (\rho \mathbf{v}) \mathrm{d}x = 0.$$

Because V is an arbitrary control volume, the equation above needs to hold for all control volumes V (especially for arbitrary small ones). It follows, that it holds pointwise:

$$\partial_t \rho + \operatorname{div}\ (\rho \mathbf{v}) = 0. \tag{2.1}$$

This is the first equation of mathematical fluid mechanics, the continuity equation.

For an *incompressible* (and homogeneous) fluid, i.e.

$$\rho(x, t) \equiv \rho_0 = const.,$$

the continuity equation reduces to

$$\operatorname{div}\ \mathbf{v} = 0,$$

which is a constraint to the velocity field $\mathbf{v}$.

Simplified spoken: If a force tries to compress an incompressible fluid, not its density, but the pressure increases. This happens for instance in water: Water is relatively heavy, so its weight by itself exerts a force that cannot be neglected. The density of water at the ground of the ocean is practically the same as at the surface, but the pressure in the depth is much higher than at the surface.

2.3 Conservation of momentum

Conservation of momentum means that the rate of change of the linear momentum equals the sum of the forces acting on a set of fluid particles or

$$\text{force} \quad = \quad \text{mass} \times \text{acceleration}.$$

Consider a fluid particle. If its position at time t is x, i.e. (x,t), then at time $t+\triangle t$ (up to the linear approximation) its position is

$$(x + \mathbf{v}\Delta t, t + \Delta t).$$

Consequently, its *acceleration* is

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \lim_{\Delta t \to 0} \frac{\mathbf{v}\left(x + \mathbf{v}\left(x,t\right)\Delta t, t + \Delta t\right) - \mathbf{v}(x,t)}{\Delta t} = \partial_t \mathbf{v} + \sum_j \mathbf{v}_j \partial_j \mathbf{v} = \partial_t \mathbf{v} + \left(\mathbf{v}\cdot\nabla\right)\mathbf{v}.$$

This derivative is called *material derivative*. The *nonlinear advective term* is in cartesian representation

$$\left(\mathbf{v}\cdot\nabla\right)\mathbf{v} = \nabla\mathbf{v}\cdot\mathbf{v} = \begin{pmatrix} \partial_1\mathbf{v}_1 & \partial_2\mathbf{v}_1 & \partial_3\mathbf{v}_1 \\ \partial_1\mathbf{v}_2 & \partial_2\mathbf{v}_2 & \partial_3\mathbf{v}_2 \\ \partial_1\mathbf{v}_3 & \partial_2\mathbf{v}_3 & \partial_3\mathbf{v}_3 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \end{pmatrix}.$$

Therefore, the product of mass and acceleration in volume V equals

$$\int_V \rho\left(\partial_t\mathbf{v} + \left(\mathbf{v}\cdot\nabla\right)\mathbf{v}\right)\mathrm{d}x,$$

which we need to balance with external (volume) forces and internal forces, which act on the volume.

External forces comprise gravity, buoyancy, Coriolis force and electromagnetic forces (in liquid metals). They are collected in a volume force term, whose net force on the volume V is given by

$$\int_V \rho f \mathrm{d}x.$$

Internal forces are forces, which the fluid exerts on itself, while it tries to get out of its own way, as for instance friction, pressure, stress or strain. Internal forces are *contact forces*: They act on the surface of the fluid element V. If σ denotes this internal force vector, then the net contribution of the internal forces equals

$$\int_{\partial V} \sigma \mathrm{d}o.$$

This yields the *momentum equation* for all control volumes V:

$$\int_V \rho\left(\partial_t\mathbf{v} + \left(\mathbf{v}\cdot\nabla\right)\mathbf{v}\right)\mathrm{d}x = \int_V \rho f \mathrm{d}x + \int_{\partial V} \sigma \mathrm{d}o. \tag{2.2}$$

To get a local, pointwise equation (for instance a partial differential equation) out of such a balance equation, the general plan is to describe the balance equation by a single volume integral over V. As this equation, again, needs to hold for arbitrary - in particular arbitrary

small - control volumes, it follows, that the equation has to hold pointwise. To enable this plan being successful, the last integral

$$\int_{\partial V} \sigma \mathrm{d}o$$

has to be replaced by a volume integral over V. For this purpose, we need more information concerning the internal forces σ.

The correct modelling of the internal forces is a crucial step on the way to predict the flow of a fluid correctly. So, we will have a closer look at this topic in the course of this chapter.

2.3.1 Conservation of angular momentum

The *angular momentum* with respect to the origin for a material volume $V = V(t)$ is defined as

$$L(V) := \int_V x \times (\rho \mathbf{v}) \mathrm{d}x$$

and its *torque* as

$$D(V) := \int_V x \times (\rho f) \mathrm{d}x + \int_{\partial V} x \times (\mathbf{n} \cdot \sigma) \mathrm{d}o.$$

The conservation law of angular momentum states, that

$$\mathrm{d}_t L(V) = D(V).$$

To simplify the notation, we introduce the *permutation tensor* $\varepsilon := (\varepsilon_{ijk})_{i,j,k=1}^3$ with elements $\varepsilon_{ijk} \in \{-1, 0, 1\}$, depending, if $\{ijk\}$ is an odd, no or an even permutation of $\{1, 2, 3\}$. Then it holds

$$(\mathbf{x} \times \mathbf{a})_i = \varepsilon_{ijk} x_j a_k \quad \forall \mathbf{x}, \mathbf{a} \in \mathbb{R}^3.$$

The transport theorem 1.4 with $\Phi = \varepsilon_{ijk} x_j \rho \mathbf{v}_k$ now implies, that

$$\mathrm{d}_t L(V) = \frac{\mathrm{d}}{\mathrm{d}t} \int_V \varepsilon_{ijk} x_j \rho \mathbf{v}_k \mathrm{d}x = \int_V \left\{ \partial_t \left(\varepsilon_{ijk} x_j \rho \mathbf{v}_k \right) + \mathrm{div} \ \left(\varepsilon_{ijk} x_j \rho \mathbf{v}_k \mathbf{v} \right) \right\} \mathrm{d}x. \tag{2.3}$$

The following identities hold:

1.
$$\partial_t \left(\varepsilon_{ijk} x_j \rho \mathbf{v}_k \right) = \varepsilon_{ijk} x_j \mathbf{v}_k \partial_t \rho + \varepsilon_{ijk} \rho x_j \partial_t \mathbf{v}_k$$

2.
$$\mathrm{div} \ \left(\varepsilon_{ijk} x_j \rho \mathbf{v}_k \mathbf{v} \right) = \partial_l \left(\varepsilon_{ijk} x_j \rho \mathbf{v}_k \mathbf{v}_l \right) = \varepsilon_{ijk} x_j \mathbf{v}_k \mathrm{div} \ (\rho \mathbf{v}) + \varepsilon_{ijk} \rho x_j \mathbf{v}_l \partial_l \mathbf{v}_k + \varepsilon_{ijk} \rho \mathbf{v}_j \mathbf{v}_k,$$

   where we used, that
$$\partial_i x_j = \delta_{ij}.$$

3.
$$\mathbf{v} \times \mathbf{v} = 0.$$

With these three identities, (2.3) can equivalently be written as

$$\mathrm{d}_t L(V) = \int_{V(t)} \{(x \times \mathbf{v})\, \partial_t \rho + (\rho x) \times \partial_t \mathbf{v} + (x \times \mathbf{v})\, \mathrm{div}\ (\rho \mathbf{v}) + (\rho x) \times (\mathbf{v} \cdot \nabla)\, \mathbf{v}\}\, \mathrm{d}x.$$

Assuming that the continuity equation

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

holds, we obtain

$$\mathrm{d}_t L(V) = \int_V x \times (\rho \partial_t \mathbf{v} + \rho\, (\mathbf{v} \cdot \nabla)\, \mathbf{v})\, \mathrm{d}x.$$

Theorem 1.3 implies that the second term in the definition of $D(V)$ can be transformed to

$$\begin{aligned}\int_{\partial V} \varepsilon_{ijk} x_j n_l \sigma_{lk} \mathrm{d}o &= \int_V \varepsilon_{ijk} \partial_l\, (x_j \sigma_{lk})\, \mathrm{d}x \\ &= \int_V \{\varepsilon_{ijk} \sigma_{jk} + \varepsilon_{ijk} x_j \partial_l \sigma_{lk}\}\, \mathrm{d}x \\ &= \int_V \{(\varepsilon : \sigma)_i + (x \times \mathrm{div}\ \sigma)_i\}\, \mathrm{d}x.\end{aligned}$$

Definition 2.2 (Frobenius (inner/scalar) product) *Let $A := (a_{ij}) \in \mathbb{R}^{m \times n}$ and $B := (b_{ij}) \in \mathbb{R}^{m \times n}$ be real matrices. The* Frobenius (inner/scalar) product *of A and B is defined as*

$$A : B := \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ij}.$$

Combining everything, we obtain

$$\int_V x \times (\rho \partial_t \mathbf{v} + \rho\, (\mathbf{v} \cdot \nabla)\, \mathbf{v})\, \mathrm{d}x = \int_V \{x \times (\rho f + \mathrm{div}\ \sigma) + \varepsilon : \sigma\}\, \mathrm{d}x.$$

Assuming that conservation of momentum

$$\rho \partial_t \mathbf{v} + \rho\, (\mathbf{v} \cdot \nabla)\, \mathbf{v} = \rho f + \mathrm{div}\ \sigma$$

holds, we obtain

$$\varepsilon : \sigma = 0.$$

It follows with the permutation tensor property of ε that

$$\sigma_{jk} - \sigma_{kj} = 0,$$

i.e., the conservation of angular momentum implies the *symmetry of the stress tensor*:

$$\sigma = \sigma^\top. \tag{2.4}$$

2.3.2 Conservation of energy

The first law of thermodynamics states that

$$\delta U = \delta Q + \delta W,$$

where U denotes the energy, W the work and Q the heat in a medium. Applied to configurations of fluid mechanics, the first law of thermodynamics leads to the postulate of a *density of internal energy* $e = e(x,t)$, such that the internal energy of a material volume $V = V(t)$ can be expressed as

$$E_{int}(V) := \int_V \rho e \mathrm{d}x.$$

Its *kinetic energy* at time t is

$$E_{kin}(V) = \frac{1}{2}\int_V \rho \left\|\mathbf{v}\right\|^2 \mathrm{d}x.$$

The temporal change in total energy

$$\mathrm{d}_t E\left(V\left(t\right)\right) = \mathrm{d}_t\left(E_{int}(V) + E_{kin}(V)\right)$$

has to be equal to the power of the acting mass forces and stresses

$$P(V) := \int_V \rho f \cdot \mathbf{v} \mathrm{d}x + \int_{\partial V} \mathbf{n} \cdot \sigma \cdot \mathbf{v} \mathrm{d}o,$$

and, additionally, the energy addition by heat sources and less the loss of energy due to heat outflow

$$Z(V) := \int_V \rho h \mathrm{d}x - \int_{\partial V} q \cdot \mathbf{n} \mathrm{d}o,$$

where $h(x,t)$ denotes the heat sources/sinks within the volume V and $q(x,t)$ the heat flux across the boundary ∂V.

Remark 2.3 1. The contribution of power is caused by translation, i.e.,

$$P = F \cdot \mathbf{v},$$

where P denotes the power, F the acting forces and $\mathbf{v}$ the velocity of the material.

2. In this energy balance we neglect the energy loss due to *radiation*.

Combining everything, we obtain the conservation equation

$$\mathrm{d}_t\left(E_{int}(V) + E_{kin}(V)\right) = P(V) + Z(V). \tag{2.5}$$

Application of the transport theorem 1.4 to $\Phi = \frac{1}{2}\rho\left\|\mathbf{v}\right\|^2$ and $\Phi = \rho e$ yields

$$\begin{aligned}
\mathrm{d}_t E_{kin}(V) &= \int_V \left\{\frac{1}{2}\partial_t\left(\rho\left\|\mathbf{v}\right\|^2\right) + \frac{1}{2}\mathrm{div}\ \left(\rho\left\|\mathbf{v}\right\|^2\mathbf{v}\right)\right\}\mathrm{d}x,\\
\mathrm{d}_t E_{int}(V) &= \int_V \left\{\partial_t\left(\rho e\right) + \mathrm{div}\ \left(\rho e \mathbf{v}\right)\right\}\mathrm{d}x.
\end{aligned}$$

Inserting these identities and the definitions of $P(V)$ and $Z(V)$ into (2.5) results in

$$\begin{aligned}
&\int_V \left\{\partial_t\left(\rho e + \frac{1}{2}\rho\left\|\mathbf{v}\right\|^2\right) + \mathrm{div}\ \left(\rho e\mathbf{v} + \frac{1}{2}\rho\left\|\mathbf{v}\right\|^2\mathbf{v}\right)\right\}\mathrm{d}x\\
=&\int_V \left\{\rho f\cdot\mathbf{v} + \rho h\right\}\mathrm{d}x + \int_{\partial V}\mathbf{n}\cdot\left(\sigma\cdot\mathbf{v} - q\right)\mathrm{d}o\\
=:&A_V + A_{\partial V}.
\end{aligned}$$

The surface integral $A_{\partial V}$ can be transformed via Theorem 1.3 - utilizing the symmetry of σ - into

$$A_{\partial V} = \int_{\partial V} n_j\left(\sigma_{ij}\mathbf{v}_i - q_j\right)\mathrm{d}o = \int_V \partial_j\left(\sigma_{ij}\mathbf{v}_i - q_j\right)\mathrm{d}x = \int_V \mathrm{div}\ \left(\sigma\cdot\mathbf{v} - q\right)\mathrm{d}x.$$

All in all, we obtain

$$\begin{aligned}&\int_V \left\{\partial_t\left(\rho e + \frac{1}{2}\rho\left\|\mathbf{v}\right\|^2\right) + \operatorname{div}\ \left(\rho e\mathbf{v} + \frac{1}{2}\rho\left\|\mathbf{v}\right\|^2\mathbf{v}\right)\right\}\mathrm{d}x\\ =&\int_V \left\{\rho f\cdot\mathbf{v} + \rho h + \operatorname{div}\ (\sigma\cdot\mathbf{v} - q)\right\}\mathrm{d}x,\end{aligned}$$

and, pointwise, the *general conservation equation of total energy (energy equation)*

$$\partial_t\left(\rho e + \frac{1}{2}\rho\left\|\mathbf{v}\right\|^2\right) + \operatorname{div}\ \left(\rho e\mathbf{v} + \frac{1}{2}\rho\left\|\mathbf{v}\right\|^2\mathbf{v}\right) = \rho f\cdot\mathbf{v} + \rho h + \operatorname{div}\ (\sigma\cdot\mathbf{v} - q)\,. \tag{2.6}$$

Therefore, the temporal and spatial change in total energy $\rho e + \frac{1}{2}\rho\left\|\mathbf{v}\right\|^2$ is determined by the external sources ρh, the increase in heat due to *mechanical* power $\rho f\cdot\mathbf{v} + \operatorname{div}\ (\sigma\cdot\mathbf{v})$ and the diffusive heat flux $\operatorname{div} q$.

Assuming, that both the continuity equation (2.1) and the conservation equation of momentum (2.2) hold, (2.6) can be reduced to

$$\partial_t(\rho e) + \operatorname{div}\ (\rho e\mathbf{v}) = \sigma : \nabla\mathbf{v} - \operatorname{div} q + \rho h.$$

2.3.3 Balance equations

We summarize the equations that we derived so far. For the conservation quantities

- ρ density of mass,
- $\rho\mathbf{v} = (\rho\mathbf{v}_i)_{i=1}^3$ momentum,
- ρE density of (total) energy,

we derived the following equations based on fundamental conservation principles:

1. Continuity equation for density ρ (conservation of mass):

$$\partial_t\rho + \operatorname{div}\ (\rho\mathbf{v}) = 0.$$

2. Momentum equation for momentum $\rho\mathbf{v}$ (conservation of momentum):

$$\partial_t(\rho\mathbf{v}) + \operatorname{div}\ (\rho\mathbf{v}\otimes\mathbf{v}) - \operatorname{div}\sigma = \rho f.$$

3. Energy equation for density of total energy $E := e + \frac{1}{2}\left\|\mathbf{v}\right\|^2$ (conservation of energy):

$$\partial_t(\rho E) + \operatorname{div}\ (\rho E\mathbf{v}) = \rho f\cdot\mathbf{v} + \operatorname{div}\ (\sigma\cdot\mathbf{v}) + \rho h - \operatorname{div} q.$$

The above variables are also called *conservative variables* and the corresponding equations *conservation equations*.

2.3.4 Kinematic properties

Exterior loads cause displacements of mass points and as a consequence geometric deformations and strains, respectively, of a mass volume. These changes in structure cause *internal counter forces* which act against the external loads and balance those in a new *resting state*. In order to describe these relations mathematically, we first define the notion of *strain*.

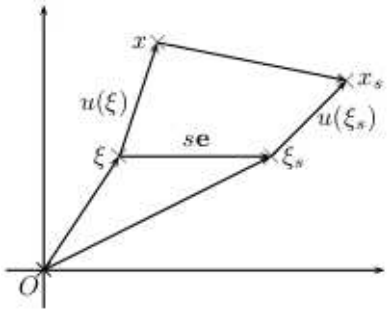

Figure 2.3: Derivation of the strain tensor of structural mechanics

2.3.4.1 Strain tensor of structural mechanics

The deformation $\xi \mapsto x = \xi + u(\xi)$, where ξ denotes the position of a material point in rest position and u the deformation of the material, causes the material volume to change its geometry. This process is called *strain*. It is our goal to describe strain in a fixed point ξ.

Consider an arbitrary directional vector $\mathbf{e} \in \mathbb{R}^3$, $\|\mathbf{e}\| = 1$, and a material line element with endpoints ξ and $\xi_s := \xi + s\mathbf{e}$. Caused by the deformation, the endpoints ξ, ξ_s pass on to

$$x = \xi + u(\xi), \quad x_s := \xi_s + u(\xi_s).$$

Then the limit

$$\mathrm{d}_e(\xi) := \lim_{s\to 0} \frac{\|x - x_s\| - \|\xi - \xi_s\|}{\|x - x_s\|}$$

describes the *relative elongation* of the material volume *in the image point* $x(\xi, t)$ *in direction* $\mathbf{e}$. With the Jacobian matrix $\nabla u(\xi)$ it holds

$$\begin{aligned}
\mathrm{d}_e(\xi) &= \lim_{s\to 0} \left\{ \frac{1}{s} \|s\mathbf{e} + u(\xi_s) - u(\xi)\| \right\} - 1 \\
&= \|\mathbf{e} + \nabla u(\xi)\mathbf{e}\| - 1 \\
&= (\mathbf{e} + \nabla u(\xi)\mathbf{e}, \mathbf{e} + \nabla u(\xi)\mathbf{e})^{\frac{1}{2}} - 1 \\
&= \left(1 + \left(\left\{\nabla u(\xi) + \nabla u(\xi)^\top + \nabla u(\xi)^\top \nabla u(\xi)\right\}\mathbf{e}, \mathbf{e}\right)\right)^{\frac{1}{2}} - 1.
\end{aligned}$$

The matrix

$$\epsilon(\xi) := \frac{1}{2}\left\{\nabla u(\xi) + \nabla u(\xi)^\top + \nabla u(\xi)^\top \nabla u(\xi)\right\}$$

with the elements

$$\epsilon_{ij}(\xi) = \frac{1}{2}\left\{\partial_j u_i + \partial_i u_j(\xi) + \sum_k \left(\partial_i u_k(\xi)\partial_j u_k(\xi)\right)\right\}$$

is called *strain tensor* (with respect to the ξ coordinate system of the rest position). The strain tensor associates every direction $\mathbf{e}$ in point ξ with a vector $\epsilon(\xi)\mathbf{e}$, which describes the strain according to

$$\mathrm{d}_e(\xi) = \left(1 + 2\left(\epsilon(\xi)\mathbf{e}, \mathbf{e}\right)\right)^{\frac{1}{2}} - 1.$$

Especially, for $\mathbf{e} = \mathbf{e}^i$ one obtains the so called *principal strains*

$$\mathrm{d}^{(i)}(\xi) = (1 + 2\epsilon_{ii}(\xi))^{\frac{1}{2}} - 1, \quad \epsilon_{ii}(\xi) := \left(\epsilon(\xi)\mathbf{e}^i, \mathbf{e}^i\right) \quad (i = 1, 2, 3).$$

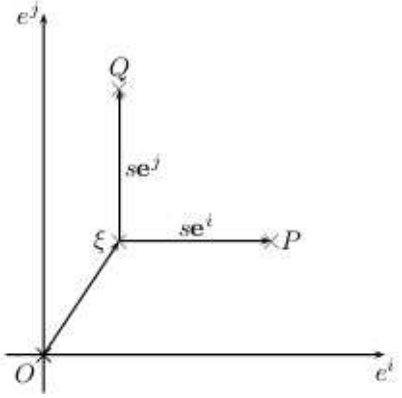

Figure 2.4: Interpretation of off-diagonal elements

To give an interpretation of the off-diagonal elements $\epsilon_{ij}(\xi)$ $(i \neq j)$, we consider two points $P := \xi + s\mathbf{e}^i$ and $Q := \xi + s\mathbf{e}^j$ in the plane spanned by $\mathbf{e}^i$ and $\mathbf{e}^j$, which form with ξ a right angle. Then the angle $\omega_{ij}(s)$ formed by the image segments $x(P) - x(\xi)$ and $x(Q) - x(\xi)$ is defined by

$$\cos\left(\omega_{ij}(s)\right) := \frac{(x(P) - x(\xi), x(Q) - x(\xi))}{\|x(P) - x(\xi)\| \, \|x(Q) - x(\xi)\|}.$$

Furthermore, it holds (exercise!)

$$\lim_{s\to 0} \frac{1}{s} \{x(P) - x(\xi)\} = (I + \nabla u(\xi))\, \mathbf{e}^i, \qquad \lim_{s\to 0} \frac{1}{s} \|x(P) - x(\xi)\| = 1 + \mathrm{d}^{(i)},$$
$$\lim_{s\to 0} \frac{1}{s} \{x(Q) - x(\xi)\} = (I + \nabla u(\xi))\, \mathbf{e}^j, \qquad \lim_{s\to 0} \frac{1}{s} \|x(Q) - x(\xi)\| = 1 + \mathrm{d}^{(j)}.$$

Consequently, one obtains for the so called *shear strain* (local change in angle) in the point $x(\xi)$ in the $\left(\mathbf{e}^i, \mathbf{e}^j\right)$ plane:

$$\cos(\omega_{ij}) = \lim_{s\to 0} \cos\left(\omega_{ij}(s)\right) = \frac{2\epsilon_{ij}}{\left(1 + \mathrm{d}^{(i)}\right)\left(1 + \mathrm{d}^{(j)}\right)}.$$

The meaning of the strain tensor's elements becomes clearer, if we restrict ourselves to very small strains, i.e., $\|\nabla u\| \ll 1$. In this case, we obtain by an approximation of first order:

$$\mathrm{d}^{(i)} \approx \epsilon_{ii} \quad (\approx 0)$$

and (by a Taylor expansion of the cosine function)

$$\omega_{ij} - \frac{\pi}{2} \approx \cos(\omega_{ij}) \approx 2\epsilon_{ij} \quad (i \neq j).$$

Next, we consider the change in volume caused by the strain. Let V_ξ be a volume fixed at the point ξ_0. Using

$$\det\left(x'(\xi)\right) = \det\left(I + \nabla u(\xi)\right),$$

we obtain:

$$\begin{aligned}
\lim_{|V_\xi|\to 0} \frac{|V_x| - |V_\xi|}{|V_\xi|} &= \lim_{|V_\xi|\to 0} \frac{\int_{V_x} \mathrm{d}x - \int_{V_\xi} \mathrm{d}\xi}{\int_{V_\xi} \mathrm{d}\xi} \\
&= \lim_{|V_\xi|\to 0} \frac{\int_{V_\xi} \det\left(x'(\xi)\right) \mathrm{d}\xi - \int_{V_\xi} \mathrm{d}\xi}{\int_{V_\xi} \mathrm{d}\xi} \\
&= \det\left(I + \nabla u(\xi_0)\right) - 1.
\end{aligned}$$

In the case of very small strains, i.e., $\|\nabla u\| \ll 1$, Taylor expansion yields

$$\det\left(I + \nabla u(\xi)\right) - 1 = \underbrace{\det(I)}_{=1} + \det'(I)\nabla u(\xi) + \mathcal{O}\left(\|\nabla u(\xi)\|^2\right) - 1.$$

The partial derivatives of the determinant of a matrix with respect to its elements are given by the cofactors of the matrix:

$$\frac{\partial}{\partial a_{ij}} \det(A) = (-1)^{i+j} co(a_{ij}), \quad i, j = 1, 2, 3,$$

where $co(a_{ij})$ is obtained as the determinant of the matrix that results if one sweeps the ith row and the jth column of A. Therefore, it holds

$$\begin{aligned}
\det\left(I + \nabla u(\xi)\right) - 1 &= \sum_{i,j} \left((-1)^{i+j} co(\delta_{ij}) \partial_j u_i(\xi)\right) + \mathcal{O}\left(\|\nabla u(\xi)\|^2\right) \\
&= \sum_{i,j} \left(\delta_{ij} \partial_j u_i(\xi)\right) + \mathcal{O}\left(\|\nabla u(\xi)\|^2\right) \\
&= \sum_{i} \left(\partial_i u_i(\xi)\right) + \mathcal{O}\left(\|\nabla u(\xi)\|^2\right) \\
&= \text{div } u(\xi) + \mathcal{O}\left(\|\nabla u(\xi)\|^2\right).
\end{aligned}$$

Since the trace of a tensor is invariant to rotations of the coordinate system, also the relative change in volume is invariant to rotations of the coordinate system (as expected).

2.3.4.2 Strain tensor of fluid mechanics

It is a typical property of fluid flow that even small acting forces result in arbitrary large displacements of the material points which is nothing else but the *flow*. In contrast to structural mechanics, not the deformation itself is of importance but the velocity at which the deformation happens.

For the specification of the term *strain velocity* we consider again an arbitrary directional vector $\mathbf{e} \in \mathbb{R}^3$, $\|e\| = 1$, and a material line element with endpoints x and $x_s := x + s\mathbf{e}$ at time t, cf. Figure 2.5. After a time interval $k = \Delta t$ the endpoints passed on to the locations

$$y = x + \mathbf{v}(x, t)k + \delta(x, k), \quad y_s = x_s + \mathbf{v}(x_s, t)k + \delta(x_s, t),$$

where the remainders are of order $\mathcal{O}\left(k^2\right)$. The relative elongation in direction $\mathbf{e}$ at time $t + k$ is then given by

$$\mathrm{d}_e(x, t + k) = \lim_{s \to 0} \frac{\|y - y_s\| - \|x - x_s\|}{\|x - x_s\|},$$

and the relative velocity of elongation at time t becomes

$$\dot{\mathrm{d}}_e(x, t) = \lim_{k \to 0} \frac{\mathrm{d}_e(x, t + k)}{k}.$$

Obviously, it holds

$$\frac{\|y - y_s\| - \|x - x_s\|}{\|x - x_s\|} = \frac{\|s\mathbf{e} - \{\mathbf{v}(x, t) - \mathbf{v}(x_s, t)\}\, k + \delta(x_s, k) - \delta(x, k)\|}{s} - 1.$$

Using the fact, that

$$\lim_{s\to 0} \frac{\|\delta(x_s,k)-\delta(x,k)\|}{s} = \mathcal{O}\left(k^2\right),$$

we obtain

$$\begin{aligned} \mathrm{d}_e(x,t+k) &= \left\|\mathbf{e}+\left\{\nabla\mathbf{v}k+\mathcal{O}\left(k^2\right)\right\}\mathbf{e}\right\|-1 \\ &= \left\{1+\left(k\left\{\nabla\mathbf{v}+\nabla\mathbf{v}^\top\right\}\mathbf{e}+\mathcal{O}\left(k^2\right)\mathbf{e},\mathbf{e}\right)\right\}^{\frac{1}{2}}-1. \end{aligned}$$

Applying the approximation

$$(1+z)^{\frac{1}{2}} = 1+\frac{z}{2}+\mathcal{O}\left(z^2\right)$$

yields

$$\dot{\mathrm{d}}_e(x,t) = \frac{1}{2}\left(\left\{\nabla\mathbf{v}+\nabla\mathbf{v}^\top\right\}\mathbf{e},\mathbf{e}\right).$$

The tensor

$$\dot{\epsilon}(x,t) := \frac{1}{2}\left\{\nabla\mathbf{v}(x,t)+\nabla\mathbf{v}^\top(x,t)\right\} \tag{2.7}$$

is called *rate of deformation tensor* (with respect to the x coordinate system which is fixed in space).

Therefore, the *velocity of elongation* with respect to direction $\mathbf{e}$ can be expressed as

$$\dot{\mathrm{d}}_e(x,t) = \left(\dot{\epsilon}(x,t)\mathbf{e},\mathbf{e}\right).$$

Especially, we obtain the *relative velocity of elongation* in direction of $\mathbf{e}^i$ as

$$\dot{\epsilon}_{ii} = \partial_i\mathbf{v}_i, \quad i=1,2,3,$$

the *relative rate of change in angle* in the $(\mathbf{e}^i,\mathbf{e}^j)$ plane as

$$\dot{\epsilon}_{ij} = \frac{1}{2}\left\{\partial_j\mathbf{v}_i+\partial_i\mathbf{v}_j\right\},$$

and, finally, the *relative rate of change in volume*

$$\mathrm{trace}(\dot{\epsilon}) = \mathrm{div}\ \mathbf{v}.$$

2.3.4.3 Optional: Derivation of rotational and strain velocities according to [7]

The gradient tensor of velocity $\nabla\mathbf{v}$ contains information about the (local) *spatial* changes of the velocity field. Consequently, it plays an important role, if we ask how infinitesimal adjacent material points are displaced in a relative view.

We denote the components of the velocity field $\mathbf{v}$ by

$$\mathbf{v} = (u,v,w)^\top.$$

We consider two points „1“ and „2“, positioned at $\mathbf{r}$ and $\mathbf{r}+\mathrm{d}\mathbf{r}$, respectively. During the time interval Δt, in linear approximation, they are shifted by $\mathbf{v}_1\Delta t$ and $\mathbf{v}_2\Delta t$, respectively. Their relative position at the time $t+\Delta t$ is thus given by the vector

$$\mathrm{d}\mathbf{r}^* = \mathrm{d}\mathbf{r}+\left(\mathbf{v}_2-\mathbf{v}_1\right)\Delta t+\mathcal{O}\left(\Delta t^2\right) \tag{2.8}$$

(see figure 2.5).

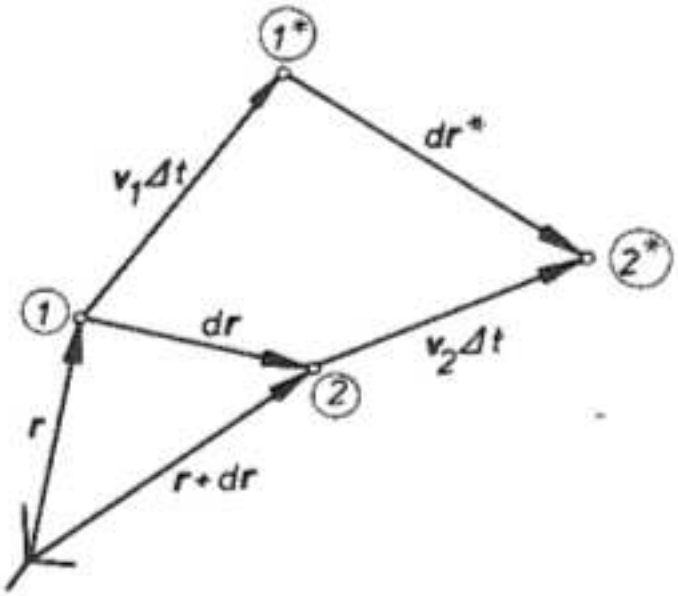

Figure 2.5: Displacement of infinitesimal adjacent points.

Now we have for the difference of the velocity components of the two infinitesimal adjacent points

$$\begin{aligned} u_2 - u_1 &= \partial_x u \mathrm{d}x + \partial_y u \mathrm{d}y + \partial_z u \mathrm{d}z \\ v_2 - v_1 &= \partial_x v \mathrm{d}x + \partial_y v \mathrm{d}y + \partial_z v \mathrm{d}z \\ w_2 - w_1 &= \partial_x w \mathrm{d}x + \partial_y w \mathrm{d}y + \partial_z w \mathrm{d}z \end{aligned}$$

or

$$\mathbf{v}_2 - \mathbf{v}_1 = \nabla \mathbf{v} \cdot \mathrm{d}\mathbf{r}.$$

Using (2.8) we achieve

$$\mathrm{d}\mathbf{r}^* - \mathrm{d}\mathbf{r} = \nabla \mathbf{v} \cdot \mathrm{d}\mathbf{r} \Delta t + \mathcal{O}\left(\Delta t^2\right).$$

The left-hand side describes the change of the distance vector to point „2“ in the time interval Δt, an observer moving with the material point „1“ sees. Dividing by Δt and then evaluating the limit for $\Delta t \to 0$ gives the following fundamental relation for infinitesimal adjacent material line elements:

$$\mathrm{d}_t\left(\mathrm{d}\mathbf{r}\right) = \nabla \mathbf{v} \cdot \mathrm{d}\mathbf{r}. \tag{2.9}$$

So, the tensor $\nabla \mathbf{v}$ transforms the relative position vector $\mathrm{d}\mathbf{r}$ into its material derivative.

We now split $\nabla \mathbf{v}$ into a symmetric part D (with $D^\top = D$) and a skew symmetric part W (with $W^\top = -W$) via

$$D := \frac{1}{2}\left(\nabla \mathbf{v} + (\nabla \mathbf{v})^\top\right), \quad W := \frac{1}{2}\left(\nabla \mathbf{v} - (\nabla \mathbf{v})^\top\right). \tag{2.10}$$

Remark 2.4 By representing D and W in cartesian coordinates, the following properties show up:

- The sum of the diagonal elements, i.e. the trace, of D corresponds to the divergence of the velocity field

$$\operatorname{div} \mathbf{v}.$$

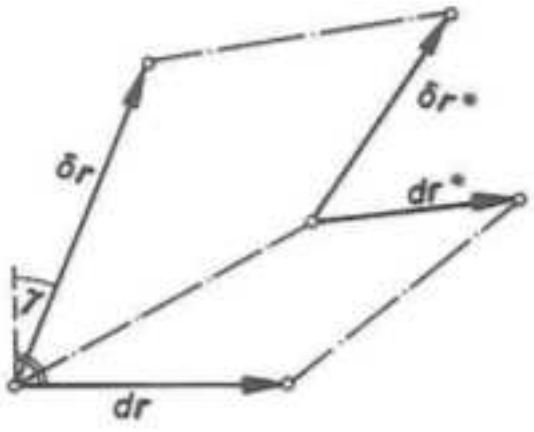

Figure 2.6: Temporal change of two material line elements.

- In W the components of the so called *vorticity vector*

$$\omega := \text{rot } \mathbf{v}$$

can be found and for an arbitrary vector $\mathbf{b}$ it holds

$$W \cdot \mathbf{b} = \frac{1}{2}\omega \times \mathbf{b}.$$

By splitting $\nabla \mathbf{v} = D + W$, the right-hand side of (2.9) decomposes into two summands, that influence the time derivative additively.

For the contribution from W it holds with remark 2.4

$$\mathrm{d}_t\,(\mathrm{d}\mathbf{r}) = \frac{1}{2}\omega \times \mathrm{d}\mathbf{r},$$

which is nothing different than a rotation of $\mathrm{d}\mathbf{r}$ with the angular velocity $\frac{\omega}{2}$. As a consequence, the tensor W is called *spin tensor*.

To interpret the tensor D we consider the change in time of the inner product of two material line elements $\mathrm{d}\mathbf{r}$ and $\delta\mathbf{r}$ with length $\mathrm{d}s$ and δs, respectively, that enclose the angle $90° - \gamma$. By

$$\mathrm{d}_t(\mathrm{d}\mathbf{r} \cdot \delta\mathbf{r}) = \mathrm{d}_t(\mathrm{d}s\delta s \sin\gamma)$$

applying the chain rule for derivatives, using (2.9), (2.10) and doing some basic transformations, one arrives at

$$\left\{ \frac{\mathrm{d}_t(\mathrm{d}s)}{\mathrm{d}s} + \frac{\mathrm{d}_t(\delta s)}{\delta s} \right\} \sin\gamma + \mathrm{d}_t\gamma \cos\gamma = 2\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s} \cdot D \cdot \frac{\delta\mathbf{r}}{\delta s}. \tag{2.11}$$

To filter a diagonal element, say D_{xx}, out of the right-hand side, we let the line elements coincide and choose their orientation (for this example) along the x-axis. On these conditions, the enclosed angle vanishes, i.e. $\gamma = 90°$, and (2.11) reduces to

$$\frac{\mathrm{d}_t(\mathrm{d}s)}{\mathrm{d}s} = D_{xx}.$$

This relation shows that the diagonal elements of the tensor D give the *elongation rate* of such line elements that currently are oriented in the direction of the basis vectors.

Analogously, it is possible to filter out an offdiagonal element out of the right-hand side of (2.11), by choosing line elements perpendicular to each other and oriented parallel to basis vectors $\mathbf{e}_i$. Then, $\gamma = 0$ and (2.11) reduces for instance to

$$\mathrm{d}_t\gamma = 2D_{xy}.$$

The offdiagonal elements therefore have half the magnitude of the velocities that describe the change of the angle between two line elements, which at the present are oriented parallel to different basis vectors. Thus, the tensor D describes the elongation rates of the edges and the rate of change of the angles between the edges of a momentarily box-shaped material volume element with an orientation parallel to the coordinate axes. Consequently, D is called *rate of deformation tensor*.

2.3.5 Material properties

2.3.5.1 Surface forces

Hydrostatic pressure:

$$\sigma_{|\mathbf{v}=0} = -pI.$$

Case of non-static fluid:

$$\sigma = -pI + \tau,$$

where τ denotes the *shear tensor*. The tensor τ depends on the deformation of the fluid: with the deformation tensor $\dot{\epsilon} = \frac{1}{2}\left(\nabla\mathbf{v} + \nabla\mathbf{v}^\top\right)$, the trial for τ reads

$$\tau = F(\dot{\epsilon}).$$

Assumptions: (material assumptions)

1. $F(0) = 0$.
2. *Symmetry:* we assume that
$$\tau = \tau^\top.$$
3. *Isotropy:* invariance with respect to orthogonal transformations, i.e.,
$$F\left(Q\dot{\epsilon}Q^\top\right) = QF\left(\dot{\epsilon}\right)Q^\top,$$
where Q is an orthogonal matrix, i.e.,
$$QQ^\top = I, \quad \det(Q) = 1.$$

Theorem 2.5 (Material tensor) *Assuming the above assumptions on the material, then*

$$F(\dot{\epsilon}) = \varphi_0 I + \varphi_1\dot{\epsilon} + \varphi_2\dot{\epsilon}^2,$$

where $\varphi_i(I_1, I_2, I_3)$ *and*

$$\begin{aligned} I_1 &= \mathrm{trace}(\dot{\epsilon}), \\ I_2 &= \frac{1}{2}\left\{\dot{\epsilon}_{ij}\dot{\epsilon}_{ji} - \dot{\epsilon}_{ii}\dot{\epsilon}_{jj}\right\}, \\ I_3 &= \det(\dot{\epsilon}). \end{aligned}$$

Proof We assume $\dot{\epsilon}$ to be a 3×3 tensor/matrix with eigenpairs

$$\{(\lambda_i, e_i)\}_{1 \le i \le 3}, \quad \lambda_i \in \mathbb{R}, \quad e_i \in \mathbb{R}^3.$$

1. We prove that $\{e_1, e_2, e_3\}$ are eigenvectors of $F(\dot{\epsilon})$.

   We consider Q the orthogonal transformation of 180° rotation around e_3.

   $$\begin{aligned} Q\dot{\epsilon} &= \dot{\epsilon}Q, \\ Q\dot{\epsilon}Q^\top &= \dot{\epsilon}. \end{aligned}$$

   Due to the isotropy condition

   $$\begin{aligned} F(\dot{\epsilon}) &= F\left(Q\dot{\epsilon}Q^\top\right) = QF(\dot{\epsilon})Q^\top, \\ F(\dot{\epsilon})Q &= QF(\dot{\epsilon}). \end{aligned}$$

   Let $W = [e_1, e_2, e_3]$.

   $$F(\mathrm{diag}(\lambda_i)) = F\left(W\dot{\epsilon}W^\top\right) = WF(\dot{\epsilon})W^\top = \mathrm{diag}(\mu_i).$$

2. We assume that the eigenvalues $\{\lambda_i\}_{1 \le i \le 3}$ are simple eigenvalues (multiplicity = 1).

   $$\det \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 \\ 1 & \lambda_3 & \lambda_3^2 \\ 1 & \lambda_3 & \lambda_3^2 \end{pmatrix} \neq 0$$

   $$\begin{aligned} \varphi_0 + \varphi_1\lambda_1 + \varphi_2\lambda_1^2 &= \mu_1(\lambda_1, \lambda_2, \lambda_3) \\ \varphi_0 + \varphi_1\lambda_2 + \varphi_2\lambda_2^2 &= \mu_2(\lambda_1, \lambda_2, \lambda_3) \\ \varphi_0 + \varphi_1\lambda_3 + \varphi_2\lambda_3^2 &= \mu_3(\lambda_1, \lambda_2, \lambda_3) \end{aligned}$$

   This linear system has a unique solution $\varphi_i = \varphi_i(\lambda_1, \lambda_2, \lambda_3) = \varphi_i(I_1, I_2, I_3)$. Due to the orthogonality of the eigensystem $\{e_1, e_2, e_3\}$ we obtain

   $$F(\dot{\epsilon}) = \varphi_0 I + \varphi_1\dot{\epsilon} + \varphi_2\dot{\epsilon}^2.$$

□

Remark 2.6 The assumption on the simplicity of the eigenvalues in step 2 can be obtained – if needed – by means of an infinitely small perturbation of $\dot{\epsilon}$.

So far, we have derived from the material assumptions, that

$$\begin{aligned} \sigma &= -pI + F(\dot{\epsilon}) \\ &= -pI + \varphi_0 I + \varphi_1\dot{\epsilon} + \varphi_2\dot{\epsilon}^2. \end{aligned}$$

If the fluid is assumed to be *Newtonian*, the term $\varphi_2\dot{\epsilon}^2$ is neglected. Consequently, the stress tensor σ becomes a *linear* relation.

Now, we switch from mathematics to physics and write the relation for a Newtonian fluid as

$$\sigma = -pI + 2\mu\dot{\epsilon} + \lambda\mathrm{tr}(\dot{\epsilon})I$$

with the *shear viscosity* μ and the *volume viscosity* λ. The shear viscosity describes how viscous a fluid is and can be measured. The volume viscosity indicates the relation between change of temperature and change of volume; as in this lecture, temperature will be considered as constant, we will not care about λ.

Now, assuming $\rho(x,t) = \rho_0$:

$$\frac{\partial u}{\partial t} - \left(\frac{\mu}{\rho_0}\right)\Delta u + (u\cdot\nabla)u + \nabla\left(\frac{p}{\rho_0}\right) = f$$
$$\nabla\cdot u = 0$$

we achieved the *unsteady incompressible Navier-Stokes equations* that provide 3+1 equations for 3+1 variables (u, p).

Remark 2.7 What is the optimal wind direction to move your sailing boat at highest speed? perpendicular to the direction of motion; cf. airplane wings: Different velocities before and behind the sail lead to different pressures leading to forces that move the boat fast. Unfortunately, this solution is unstable.

2.3.5.2 Reynolds number

We assume that the flow occurs in a domain Ω and the flow geometry has some characteristic length L. For example, L can be the length of an airplane.
Thus, we can geometrically scale the domain to get a rescaled domain $\Omega^* = \frac{\Omega}{L}$. Ω and Ω^* are geometrically similar.

$$x^* = \frac{x}{L}, \qquad L: \text{ reference length}$$
$$u^* = \frac{u}{V}, \qquad V: \text{ reference speed}$$
$$t^* = V\cdot\frac{t}{L},$$
$$p^* = \frac{p}{\rho_0 V^2}$$

V is a reference speed, for example the free stream speed away from the model airplane.
If we denote partial derivatives with respect to the rescaled dimensionless variables by $\frac{\partial}{\partial x^*}$, then the chain rule implies

$$\frac{\partial u^*}{\partial x^*} = \frac{L}{V}\frac{\partial u}{\partial x} \quad \text{for example.}$$

The Navier-Stokes equations for incompressible flows in the rescaled variables becomes

$$\frac{\partial u^*}{\partial t^*} - \left(\frac{\mu}{\rho_0 V L}\right)\underbrace{\Delta^* u^*}_{\text{viscous forces}} + \underbrace{(u^*\cdot\nabla^*)u^*}_{\text{inertial forces}} + \nabla^* p^* = f$$
$$\nabla^* u^* = 0.$$

Definition 2.8 (Dynamic similarity) *Flows in similar geometry Ω and $\Omega^* = \frac{\Omega}{L}$ are dynamically similar, if the parameters of the flow are such that the quantities of the flows*

$$Re = \frac{\rho_0 V L}{\mu} \tag{2.12}$$

coincide.
The dimensionless parameter Re is called the Reynolds number*.*

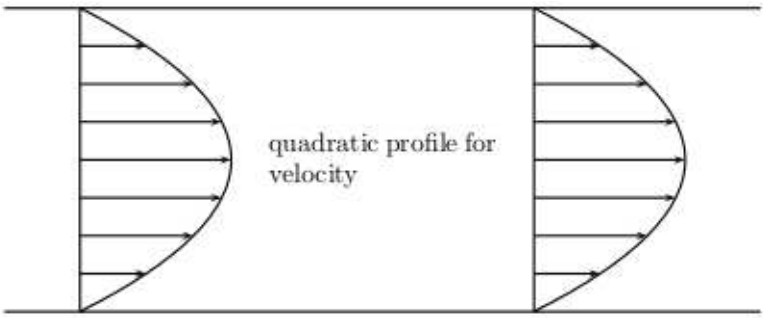

Figure 2.7: Poiseuille flow.

Remark 2.9 A step function (eg. deriving from an earth quake) will remain unchanged during time passing by, as long as the viscous forces are neglected. But taking the viscous forces into consideration leads to a smoothing of the function after some time.

Remark 2.10 The dynamic similarity can be very usefull: To verify for instance your simulation results of an airplane, instead of building the airplane in full size, you can build a smaller model and adapt the other parameters (increase V or decrease μ) in such a way that Re is kept constant.

$$\mathrm{Re} = \frac{\text{inertial forces}}{\text{viscous forces}}$$

If Re is close to 0, the viscous forces dominate inertial forces, as occurs in a highly viscous fluid moving slowly.
For very large Reynolds numbers the viscous forces can be neglected. This often occurs in flows of gases.

Examples:

- model airplane, 1 ms^{-1}: $Re \sim 10^4$
- car, 3 ms^{-1}: $Re \sim 10^5$
- small airplane, 30ms^{-1}: $Re \sim 10^7$

2.3.5.3 Poiseuille flow

At low velocities, one has a laminar flow. By increasing the velocity, eventually one reaches the critical Reynolds number where turbulences occur and the flow gets chaotic.

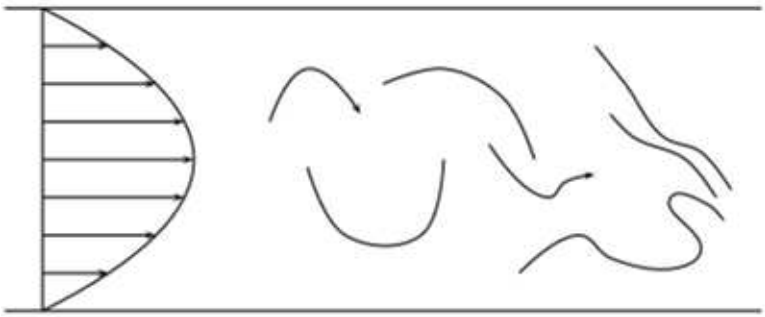

Figure 2.8: Transition to a chaotic flow.

3 Approximating Steady Flows

This chapter considers the problem of approximating a steady solution $(\mathbf{v}, p)$ of the Navier-Stokes equations (NSEs):

$$\begin{cases} -\nu\Delta\mathbf{v} + (\mathbf{v}\cdot\nabla)\,\mathbf{v} + \nabla p &= f, \quad \text{in } \Omega \\ \operatorname{div}\mathbf{v} &= 0, \quad \text{in } \Omega \\ \mathbf{v} &= 0, \quad \text{on } \partial\Omega \\ \int_\Omega p\mathrm{d}x &= 0. \end{cases} \tag{3.1}$$

The steady NSE present the two interesting problems of stability bounds for the pressure (coupling between $\operatorname{div}\mathbf{v} = 0$ and ∇p), which, mathematically, is equivalent to a *saddle-point problem*, and stability bounds for the velocity (due to the nonlinearity). In the methods we consider, boundedness of the pressure by problem data is ensured by using elements which satisfy the so called *discrete* $\inf-\sup$ *condition*. It is a discrete analog to the $\inf-\sup$ stability condition which is needed in the (weak) continuous theory to ensure boundedness of the continuous pressure. The other main difficulty is the nonlinearity. The solution of this problem which ensures the physical energy bounds hold for approximate solutions is to explicitly skew-symmetrize the nonlinearity in the equations for the approximate solution.

3.1 The Stokes problem: Introduction to Mixed Methods

In this section we study the *Stokes problem* which is a simplification of NSE. It is the „simplest“ subproblem of the full NSE where the conditions for the stability bounds for the pressure can be established. As we will see later, the solution to this problem in the case of the Stokes problem is also a solution in the case of the nonlinear NSE.

The Stokes problem consists of finding the fluid velocity $\mathbf{v} : \Omega \to \mathbb{R}^d$ and pressure $p : \Omega \to \mathbb{R}$ defined in the flow domain $\Omega \subset \mathbb{R}^d$ satisfying

$$\begin{cases} -\Delta\mathbf{v} + \nabla p &= f, \quad \text{in } \Omega \\ \operatorname{div}\mathbf{v} &= 0, \quad \text{in } \Omega \\ \mathbf{v} &= 0, \quad \text{on } \partial\Omega \end{cases} \tag{3.2}$$

Note that ∇p occurs in the Stokes problem rather than p and that there is (under the above most common boundary conditions at least) no pressure boundary condition. Thus, the pressure can, at best, be determined only up to an additive constant

$$\nabla(p + C) = \nabla p.$$

For this reason the pressure is normalized in some way to determine the arbitrary additive constant; the most mathematically convenient way is by

$$\int_\Omega p\mathrm{d}x = 0.$$

First, we derive a variational formulation of (3.2). Let $(\mathbf{v}, p)$ be a classical solution of the Stokes problem. Multiply (3.2) by (smooth enough) functions (φ, q) and integrate:

$$\begin{cases} \int_\Omega (-\Delta \mathbf{v} \cdot \varphi + \nabla p \cdot \varphi)\, \mathrm{d}x & = \int_\Omega f \cdot \varphi \mathrm{d}x, \\ \int_\Omega \operatorname{div} \mathbf{v} q \mathrm{d}x & = 0. \end{cases}$$

Applying the divergence theorem 1.3, term by term, gives

$$-\int_\Omega \Delta \mathbf{v} \cdot \varphi \mathrm{d}x = \int_\Omega \nabla \mathbf{v} : \nabla \varphi \mathrm{d}x - \int_{\partial\Omega} \nabla \mathbf{v} \cdot \varphi \cdot \mathbf{n} \mathrm{d}o.$$

If φ vanishes on $\partial\Omega$ we thus have

$$-\int_\Omega \Delta \mathbf{v} \cdot \varphi \mathrm{d}x = \int_\Omega \nabla \mathbf{v} : \nabla \varphi \mathrm{d}x.$$

Similarly, if φ again vanishes on $\partial\Omega$,

$$\int_\Omega \nabla p \cdot \varphi \mathrm{d}x = -\int_\Omega p \operatorname{div} \varphi \mathrm{d}x.$$

Thus, for all φ vanishing on $\partial\Omega$ and smooth enough and all q smooth enough, $(\mathbf{v}, p)$ satisfies

$$\begin{cases} \int_\Omega (\nabla \mathbf{v} : \nabla \varphi - p \operatorname{div} \varphi)\, \mathrm{d}x & = \int_\Omega f \cdot \varphi \mathrm{d}x \\ \int_\Omega \operatorname{div} \mathbf{v} q \mathrm{d}x & = 0. \end{cases} \tag{3.3}$$

The last system of equations is still well-defined in the following function spaces: Define the velocity space X as

$$X := \left(H_0^1(\Omega)\right)^d = \left(\left\{v \in C^1(\Omega) :\ v \in L^2(\Omega),\ \partial_i v \in L^2(\Omega) \text{ ex. in weak sense } \forall i,\ v = 0 \text{ on } \partial\Omega\right\}\right)^d.$$

The pressure space Q does not require any differentiability since no derivatives of p or q appear in (3.3). On the other hand, it must account for the fact that if $(\mathbf{v}, p)$ is a solution of (3.2), then so is $(\mathbf{v}, p + C)$ for any constant C. Accordingly, to fix the value of the undetermined, additive constant it is usual to impose the condition of mean value zero and thus define

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

We thus come to the variational formulation of (3.2):

$$\text{Find } \mathbf{v} \in X,\ p \in Q \text{ satisfying} \quad \begin{cases} (\nabla \mathbf{v}, \nabla \varphi) - (p, \operatorname{div} \varphi) & = (f, \varphi) \quad \forall \varphi \in X \\ (\operatorname{div} \mathbf{v}, q) & = 0 \quad \forall q \in Q \end{cases}. \tag{3.4}$$

Lemma 3.1 (Continuity)

$$b(\mathbf{v}, q) := \int_\Omega \operatorname{div} \mathbf{v} q \mathrm{d}x$$

is continuous on $X \times Q$:

$$|b(\mathbf{v}, q)| \le C \|q\|_Q \|\mathbf{v}\|_X.$$

As a consequence the divergence free subspace *V of X*

$$V := \{v \in X :\quad (\operatorname{div} v, q) = 0\ \forall q \in Q\}$$

is a closed subspace of X.

Proof The inequality is obvious due to Hölder inequality and definitions of the norms. Since $b(\mathbf{v}, q)$ is bilinear it follows equally easily that V is a subspace. The key to showing V is closed is to use continuity. By continuity, let $\mathbf{v}_n \in V$ be a sequence such that $\mathbf{v}_n \to \mathbf{v}$ in X and let $q \in Q$ be fixed: then $b(\mathbf{v}_n, q) \to b(\mathbf{v}, q)$ for $n \to \infty$. Since $b(\mathbf{v}_n, q) \equiv 0$ it follows from continuity that $b(\mathbf{v}, q) = 0$, i.e., $v \in V$. □

Definition 3.2 *For $f \in L^2(\Omega)$, the H^{-1} norm and the V^* norm of f are*

$$\|f\|_{-1} := \sup_{v \in X} \frac{|(f, v)|}{\|\nabla v\|}, \quad \|f\|_* := \sup_{v \in V} \frac{|(f, v)|}{\|\nabla v\|}.$$

Definition 3.3 *The function spaces $H^{-1}(\Omega)$ and V^* are, respectively, the closures of $L^2(\Omega)$ in, respectively, $\|\cdot\|_{-1}$ and $\|\cdot\|_*$.*

We don't want to go to far afield into the question of existence of solutions to the Stokes problem in various function Sobolev spaces with data in other Sobolev spaces. Hoewever, there is one very important theoretical result (proved by Ladyzhenskaya), the *continuous* $\inf - \sup$ *condition*, which we give here for future reference.

Proposition 3.4 (The continuous $\inf - \sup$ condition) *There is a constant $\beta > 0$ such that*

$$\inf_{q \in Q} \sup_{v \in X} \frac{(\text{div } v, q)}{\|\nabla v\| \, \|q\|} \geq \beta > 0. \tag{3.5}$$

To indicate the critical importance of the last lemma and proposition, note that the lemma implies V is in fact also a Hilbert space. Since div $\mathbf{v} = 0$ (weakly) the solution of the Stokes problem lies in V. It has thus the following formulation in V:

$$\text{Find } \mathbf{v} \in V \text{ satisfying} \quad (\nabla \mathbf{v}, \nabla \varphi) = (f, \varphi) \ \ \forall \varphi \in V. \tag{3.6}$$

In this formulation, existence and uniqueness of $\mathbf{v}$ follow from the Lax-Milgram theorem, as does the simple bound on the fluid velocity:

$$\|\nabla \mathbf{v}\| \leq \|f\|_* .$$

This bound shows that the velocity is bounded by the body force, the most fundamental of many different types of stability. Thus, Lemma 3.1 and the Lax-Milgram theorem immediately imply existence of a velocity satisfying the variational formulation of the Stokes problem in V. (The $\inf - \sup$ condition plays the key role of ensuring that, given the unique velocity, there is a corresponding pressure.)

Proposition 3.5 (Existence) *For any $f \in L^2(\Omega)$, there exists a unique velocity in V solving the Stokes problem. This velocity satisfies the a-priori bound*

$$\|\nabla \mathbf{v}\| \leq \|f\|_* .$$

The $\inf - \sup$ condition is also critical to bounding the fluid pressure, i.e., showing the pressure is stable in a fundamental sense. To see this, note that the $\inf - \sup$ condition (3.5) is equivalent to

$$\sup_{0 \neq v \in X} \frac{(\text{div } v, q)}{\|\nabla v\|} \geq \beta \, \|q\| \quad \text{for any } q \in Q.$$

To use this, our strategy is to isolate the $(\operatorname{div}\varphi, p)$ term and seek an upper bound resembling

$$(\operatorname{div}\varphi, p) = \text{everything else} \leq \cdots \leq \{\text{terms}\}\,\|\nabla\varphi\|\,.$$

Next, we divide both sides by $\|\nabla\varphi\|$ and take a supremum over $v \in X$. To be specific, rearranging (3.4)

$$\begin{aligned}(p, \operatorname{div}\varphi) &= (\nabla\mathbf{v}, \nabla\varphi) - (f, \varphi)\\ &\leq \|\nabla\mathbf{v}\|\,\|\nabla\varphi\| + \|f\|_*\,\|\nabla\varphi\|\\ &\leq (\|\nabla\mathbf{v}\| + \|f\|_*)\,\|\nabla\varphi\|\,.\end{aligned}$$

Thus,

$$\beta\,\|p\| \leq \sup_{0\neq\varphi\in X} \frac{(\operatorname{div}\varphi, p)}{\|\nabla\varphi\|} \leq (\|\nabla\mathbf{v}\| + \|f\|_*)\,.$$

The upper bound $\|\nabla\mathbf{v}\| \leq \|f\|_*$ gives

$$\|p\| \leq \frac{2\,\|f\|_*}{\beta}.$$

Thus, adding the bounds for the velocity and pressure together gives

$$\|\nabla\mathbf{v}\| + \|p\| \leq \left(1 + \frac{2}{\beta}\right)\|f\|_*\,,$$

proving a stability bound on the fluid velocity and pressure.

Definition 3.6 (The discrete $\inf-\sup$ condition) *The finite-dimensional spaces $X_h \subset X$ and $Q_h \subset Q$ fulfil the* discrete $\inf-\sup$ condition*, if there is a constant $\beta_h > 0$ such that*

$$\inf_{q_h\in Q_h}\sup_{v_h\in X_h} \frac{(\operatorname{div} v_h, q_h)}{\|\nabla v_h\|\,\|q_h\|} \geq \beta_h > 0, \tag{3.7}$$

where β_h is bounded away from zero uniformly in h.

Remark 3.7 (Interpretation of discrete $\inf-\sup$ condition) If (3.4) is discretized with finite elements, then the system matrix of the discrete system has the following block representation:

$$\begin{pmatrix} A & -B \\ B^\top & 0 \end{pmatrix},$$

where

$$\begin{aligned}A &:= (a_{ij})_{i,j},\ a_{ij} := (\nabla\varphi_j, \nabla\varphi_i)\,,\\ B &:= (b_{ij})_{i,j},\ b_{ij} := (q_j, \operatorname{div}\varphi_i)\,.\end{aligned}$$

For the existence and uniqueness of a discrete solution it is both a necessary and sufficient condition that the above matrix is regular, i.e., that the discrete linear system has a unique solution. The critical part is the „0“ block: if this part of the matrix is too big, then the linear system is under-determined, i.e., the rows are linearly dependent, and the solution is not unique anymore. If, on the other hand, this block is too small, then there are posed too many conditions on the velocity and the system contains contradicting rows, i.e., the linear system is over-determined, and there doesn't exist any solution. At this point, the discrete $\inf-\sup$ condition guarantees exactly that the „0“ block has precisely the right size to obtain a regular linear system!

3.2 Formulation and Stability of the Approximation

We now turn our attention to the NSE again. Analogously to the previous section about the Stokes problem, we can derive a weak formulation of the NSE. It reads:

$$\text{Find } \mathbf{v} \in X,\ p \in Q \text{ satisfying } \quad \begin{cases} \nu\left(\nabla \mathbf{v}, \nabla \varphi\right) + \left(\left(\mathbf{v} \cdot \nabla\right)\mathbf{v}, \varphi\right) - \left(p, \operatorname{div} \varphi\right) & = (f, \varphi) \quad \forall \varphi \in X \\ \left(\operatorname{div} \mathbf{v}, q\right) & = 0 \quad \forall q \in Q \end{cases}, \tag{3.8}$$

with the spaces X and Q defined as above.

The solution of the stated first problem — stability of the pressure — is to use, as for the Stokes problem, finite element spaces

$$X_h \subset X, \quad Q_h \subset Q$$

satisfying (3.7).

The second problem is treated by carefully formulating the nonlinearity in the discrete problem. To this end, we first state some properties of the trilinear form that is associated with the nonlinearity.

Lemma 3.8 (Skew-symmetry) *If* $\mathbf{v}, \nabla \mathbf{v} \in L^2(\Omega)$, $\operatorname{div} \mathbf{v} = 0$ *and* $\mathbf{v} \cdot \mathbf{n} = 0$ *on* $\partial\Omega$, *then it holds*

$$\int_\Omega \mathbf{v} \cdot \nabla \mathbf{v} \cdot \mathbf{v} \mathrm{d}x = 0. \tag{3.9}$$

More generally,

$$\int_\Omega \mathbf{u} \cdot \nabla \mathbf{v} \cdot \mathbf{w} \mathrm{d}x = -\int_\Omega \mathbf{u} \cdot \nabla \mathbf{w} \cdot \mathbf{v} \mathrm{d}x$$

for any such $\mathbf{u}, \mathbf{v}, \mathbf{w}$.

Proof Cf. [9], Chapter 6, Lemma 12. □

Remark 3.9 Especially, (3.9) holds for all $\mathbf{v} \in V$ because $\mathbf{v} \cdot \mathbf{n}$ vanishes *exactly* on the boundary and div $\mathbf{v}$ is *exactly* zero.

Lemma 3.10 (Continuity of the trilinear form) *There is a finite constant* $M = M(\Omega)$ *such that for all* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in X$,

$$\left|\left((\mathbf{u} \cdot \nabla)\mathbf{v}, \mathbf{w}\right)\right| \leq M \left\|\nabla \mathbf{u}\right\| \left\|\nabla \mathbf{v}\right\| \left\|\nabla \mathbf{w}\right\|. \tag{3.10}$$

Proof Cf. [9], Chapter 6, Lemma 13. □

Remark 3.11 Since V is a closed subspace of X, the trilinear form is also continuous on $V \times V \times V$.

Difficulties can arise because an approximate solution

$$\mathbf{v}_h \in V_h := \{\mathbf{v} \in X_h : \quad (q_h, \operatorname{div} \mathbf{v}) = 0 \ \forall q_h \in Q_h\}$$

is approximately but (except in very special cases) never exactly divergence free:

$$\operatorname{div} \mathbf{v}_h \neq 0 \quad \text{and} \quad V_h \nsubseteq V \quad \text{in general!}$$

To formulate the discrete problem so as to eliminate any such potential difficulties, define the (explicitly skew symmetrized) trilinear form

$$b^*(\mathbf{u},\mathbf{v},\mathbf{w}) := \frac{1}{2}\left((\mathbf{u}\cdot\nabla)\mathbf{v},\mathbf{w}\right) - \frac{1}{2}\left((\mathbf{u}\cdot\nabla)\mathbf{w},\mathbf{v}\right).$$

We have associated with the nonlinearity the finite continuity constants (which depend on the domain Ω):

$$M = M(\Omega) := \sup_{\mathbf{u},\mathbf{v},\mathbf{w}\in X} \frac{|((\mathbf{u}\cdot\nabla)\mathbf{v},\mathbf{w})|}{\|\nabla\mathbf{u}\|\,\|\nabla\mathbf{v}\|\,\|\nabla\mathbf{w}\|} < \infty,$$
$$N = N(\Omega) := \sup_{\mathbf{u},\mathbf{v},\mathbf{w}\in V} \frac{|((\mathbf{u}\cdot\nabla)\mathbf{v},\mathbf{w})|}{\|\nabla\mathbf{u}\|\,\|\nabla\mathbf{v}\|\,\|\nabla\mathbf{w}\|} < \infty.$$

Since $V \subset X$ it follows that $N \le M$.

Lemma 3.12 (Skew-symmetry and continuity) *For $\mathbf{u} \in V$, $\mathbf{v},\mathbf{w} \in X$,*

$$((\mathbf{u}\cdot\nabla)\mathbf{v},\mathbf{w}) = b^*(\mathbf{u},\mathbf{v},\mathbf{w}).$$

Further,

$$b^*(\mathbf{u},\mathbf{v},\mathbf{v}) = 0 \quad \textit{for any} \quad \mathbf{u},\mathbf{v} \in X,$$

and

$$|b^*(\mathbf{u},\mathbf{v},\mathbf{w})| \le M\,\|\nabla\mathbf{u}\|\,\|\nabla\mathbf{v}\|\,\|\nabla\mathbf{w}\| \quad \forall \mathbf{u},\mathbf{v},\mathbf{w} \in X,$$

for the same $M = M(\Omega)$.

Proof Exercise. □

Thus, another variational formulation of a solution of (3.8) which is equivalent for the continuous problem is as follows:

$$\text{Find } \mathbf{v}\in X,\ p\in Q \text{ satisfying} \quad \begin{cases} \nu\,(\nabla\mathbf{v},\nabla\varphi) + b^*(\mathbf{v},\mathbf{v},\varphi) - (p,\operatorname{div}\varphi) &= (f,\varphi) \quad \forall\varphi\in X \\ (\operatorname{div}\mathbf{v},q) &= 0 \quad \forall q\in Q \end{cases}. \tag{3.11}$$

The corresponding finite element approximation that we consider reads: Find $\mathbf{v}_h \in X_h$, $p_h \in Q_h$ satisfying

$$\begin{cases} \nu\,(\nabla\mathbf{v}_h,\nabla\varphi_h) + b^*(\mathbf{v}_h,\mathbf{v}_h,\varphi_h) - (p_h,\operatorname{div}\varphi_h) &= (f,\varphi_h) \quad \forall\varphi_h\in X_h \\ (\operatorname{div}\mathbf{v}_h,q_h) &= 0 \quad \forall q_h\in Q_h \end{cases}. \tag{3.12}$$

The formulation (3.12) can be written as follows: Find $\mathbf{v}_h \in V_h$ satisfying

$$\nu\,(\nabla\mathbf{v}_h,\nabla\varphi_h) + b^*(\mathbf{v}_h,\mathbf{v}_h,\varphi_h) = (f,\varphi_h) \quad \forall\varphi_h\in V_h. \tag{3.13}$$

Lemma 3.13 (Stability) *The finite element approximation* (3.12) *is stable,*

$$\nu\,\|\nabla\mathbf{v}_h\| \le \|f\|_{-1},$$

and if (3.7) *holds, then*

$$\|p_h\| \le \frac{1}{\beta}\left(2 + \frac{M}{\nu^2}\,\|f\|_{-1}\right)\|f\|_{-1}.$$

Proof For the first result, set $\varphi_h = \mathbf{v}_h$ in (3.13). This is equivalent to setting $\varphi_h = \mathbf{v}_h$ and $q_h = p_h$ in (3.12) and adding. Using $b^*(\mathbf{v}_h, \mathbf{v}_h, \mathbf{v}_h) = 0$ we get

$$\nu \|\nabla \mathbf{v}_h\|^2 = (f, \mathbf{v}_h) \leq \|f\|_{-1} \|\nabla \mathbf{v}_h\|,$$

giving the first inequality. For the stability bound of the pressure, solve (3.12) for $(p_h, \text{div } \varphi_h)$:

$$(p_h, \text{div } \varphi_h) = -(f, \varphi_h) + \nu (\nabla \mathbf{v}_h, \nabla \varphi_h) + b^*(\mathbf{v}_h, \mathbf{v}_h, \varphi_h),$$

so

$$(p_h, \text{div } \varphi_h) \leq \|f\|_* \|\nabla \varphi_h\| + \nu \|\nabla \mathbf{v}_h\| \|\nabla \varphi_h\| + M \|\nabla \mathbf{v}_h\| \|\nabla \mathbf{v}_h\| \|\nabla \varphi_h\|.$$

Thus,

$$\frac{(p_h, \text{div } \varphi_h)}{\|\nabla \varphi_h\|} \leq \|f\|_* + \nu \|\nabla \mathbf{v}_h\| + M \|\nabla \mathbf{v}_h\|^2.$$

Taking the supremum over $\varphi_h \in X_h$ and using $\nu \|\nabla \mathbf{v}_h\| \leq \|f\|_{-1}$ and the assumed condition (3.7) gives the pressure bound. □

It is useful to think of the error analysis as a combination of three ideas:

- The Céa's lemma shows how to handle the contribution of the viscous terms.
- The inf − sup condition shows how to handle the coupling between div $\mathbf{v} = 0$ and ∇p in the error analysis.
- The small data condition
$$\frac{M}{\nu^2} \|f\|_{-1} \leq \alpha < 1$$
will be used to control the nonlinearity in the error analysis.

The error analysis is complex but the underlying ideas are simple: To give some understanding how to handle the nonlinear terms we first consider an example with contractive properties.

Suppose $T : X \to X$ satisfies

$$\|T(x) - T(y)\|_X \leq \alpha \|x - y\|_X \quad \forall x, y \in X$$

for some $\alpha < 1$. Let $X_h \subset X$ be a subspace. Let x^* be the unique fixed point of T:

$$x^* - T(x^*) = 0.$$

Let $x_h \in X_h \subset X$ be the Galerkin approximation of x^*

$$(x_h - T(x_h), y_h) = 0 \quad \forall y_h \in X_h.$$

Let us consider the error in x^*:

$$((x^* - x_h) - (T(x^*) - T(x_h)), y_h)_X = 0.$$

Thus,

$$((x^* - x_h), y_h)_X = ((T(x^*) - T(x_h)), y_h)_X.$$

Let $P_h : X \to X_h$ denote the orthogonal projection into X_h. Set $y_h = P_h(x^* - x_h)$. Then,

$$\|P_h(x^* - x_h)\|_X^2 \leq \alpha \|x^* - x_h\|_X \|P_h(x^* - x_h)\|_X.$$

Thus,

$$\|P_h(x^* - x_h)\|_X \leq \alpha \|x^* - x_h\|_X.$$

Furthermore,

$$\begin{aligned} \|x^* - x_h\|_X^2 &\leq \|x^* - P_h x^*\|_X^2 + \|P_h(x^* - x_h)\|_X^2 \\ &\leq \|x^* - P_h x^*\|_X^2 + \alpha^2 \|x^* - x_h\|_X^2. \end{aligned}$$

Theorem 3.14 (Error in Galerkin approximation of fixed points) *Let x^* be the unique fixed point of a contractive map in X with contraction constant α and let x_h denote its Galerkin approximation. Then, the error satisfies:*

$$\|x^* - x_h\|_X \leq \left(1 - \alpha^2\right)^{-\frac{1}{2}} \inf_{w_h \in X_h} \|x^* - w_h\|_X .$$

Theorem 3.15 *Suppose the small data condition*

$$\frac{N_h}{\nu^2} \|f\|_h \leq \alpha < 1,$$

where

$$N_h := \sup_{\mathbf{u}_h, \mathbf{v}_h, \mathbf{w}_h \in V_h} \frac{b^*(\mathbf{u}_h, \mathbf{v}_h, \mathbf{w}_h)}{\|\nabla \mathbf{u}_h\| \, \|\nabla \mathbf{v}_h\| \, \|\nabla \mathbf{w}_h\|},$$

holds. Then there is at most one solution for (3.13).

Proof First, we derive the error equation. This is the nonlinear equivalent of the Galerkin orthogonality.

$$\forall \varphi_h \in X_h, \quad \forall q_h \in Q_h :$$

$$\begin{cases} \nu \left(\nabla(\mathbf{v} - \mathbf{v}_h), \nabla \varphi_h\right) + b^*(\mathbf{v}, \mathbf{v}, \varphi_h) - b^*(\mathbf{v}_h, \mathbf{v}_h, \varphi_h) - (p, \text{div } \varphi_h) & = 0 \\ \text{div } \mathbf{v} & = 0 \\ (\text{div } \mathbf{v}_h, q_h) & = 0 \end{cases}$$

It's an important refinement of this equation to note that since $\mathbf{v}_h \in X_h$, $(\text{div } \mathbf{v}_h, q_h) = 0$, we can write:

$$(p, \text{div } \varphi_h) = (p - q_h, \text{div } \varphi_h) \quad \forall q_h \in Q_h, \ \varphi_h \in X_h.$$

The error equation then becomes

$$\nu \left(\nabla(\mathbf{v} - \mathbf{v}_h), \nabla \varphi_h\right) + b^*(\mathbf{v}, \mathbf{v}, \varphi_h) - b^*(\mathbf{v}_h, \mathbf{v}_h, \varphi_h) - (p - q_h, \text{div } \varphi_h) = 0 \quad \forall q_h \in Q_h, \ \varphi_h \in X_h.$$

1. Write $\mathbf{v} - \mathbf{v}_h = (\mathbf{v} - \varphi_h) - (\mathbf{v}_h - \varphi_h) = \eta - \phi_h$ (where $\phi_h \in X_h$), where φ_h is an optimal approximation of $\mathbf{v}$ in X_h.
2. Put the ϕ_h terms on one side and the η term on the other, set $\varphi_h = \phi_h$ and get a bound of $\|\nabla \phi_h\|$ in terms of $\|\nabla \eta\|$.
3. Quadratic nonlinearities are usually treated in the following way:

$$aa - bb = a(a - b) + (a - b)b.$$

4. Apply Cauchy-Schwartz inequalities to the right-hand side for η and ϕ_h. Use the small data condition to hide the ϕ_h terms.
5. Apply the triangle inequality $\|\nabla(\mathbf{v} - \mathbf{v}_h)\| \leq \|\nabla \eta\| + \|\nabla \phi_h\|$.
6. Take the infimum over $\varphi_h \in X_h$.

Step 1 gives

$$\mathbf{v} - \mathbf{v}_h = \eta - \phi_h.$$

We obtain

$$\nu \left(\nabla \phi_h, \nabla \varphi_h\right) = \frac{1}{Re} \left(\nabla \eta, \nabla \varphi_h\right) + b^*(\mathbf{v}, \mathbf{v}, \varphi_h) - b^*(\mathbf{v}_h, \mathbf{v}_h, \varphi_h) - (p - q_h, \text{div } \varphi_h) \quad \forall \varphi_h \in X_h, \ \forall q_h \in Q_h.$$

Step 2:

$$\nu \|\nabla \phi_h\|^2 = \frac{1}{Re} (\nabla \eta, \nabla \phi_h) + b^*(\mathbf{v}, \mathbf{v}, \phi_h) - b^*(\mathbf{v}_h, \mathbf{v}_h, \phi_h) - (p - q_h, \text{div } \phi_h) \quad \forall q_h \in Q_h. \tag{3.14}$$

Step 3:

$$\begin{aligned} b^*(\mathbf{v}, \mathbf{v}, \phi_h) - b^*(\mathbf{v}_h, \mathbf{v}_h, \phi_h) &= b^*(\mathbf{v}, \mathbf{v} - \mathbf{v}_h, \phi_h) + b^*(\mathbf{v} - \mathbf{v}_h, \mathbf{v}_h, \phi_h) \\ &= b^*(\mathbf{v}, \eta - \phi_h, \phi_h) + b^*(\eta - \phi_h, \mathbf{v}_h, \phi_h) \\ &= b^*(\mathbf{v}, \eta, \phi_h) + b^*(\eta, \mathbf{v}_h, \phi_h) - b^*(\phi_h, \mathbf{v}_h, \phi_h). \end{aligned}$$

It follows

$$\begin{aligned} |b^*(\mathbf{v}, \mathbf{v}, \phi_h) - b^*(\mathbf{v}_h, \mathbf{v}_h, \phi_h)| \leq\ & M \|\nabla \mathbf{v}\| \|\nabla \eta\| \|\nabla \phi_h\| \\ & + M \|\nabla \mathbf{v}_h\| \|\nabla \eta\| \|\nabla \phi_h\| \\ & + M \|\nabla \mathbf{v}_h\| \|\nabla \phi_h\|^2 . \end{aligned}$$

Inserting this bound in (3.14) gives

$$(\nu - M \|\nabla \mathbf{v}_h\|) \|\nabla \phi_h\|^2 \leq \nu (\nabla \eta, \nabla \phi_h) - (p - q_h, \text{div } \phi_h) + M (\|\mathbf{v}\| + \|\mathbf{v}_h\|) \|\nabla \eta\| \|\nabla \phi_h\| .$$

Step 4: Now we use the idea of step 4 ($ab \leq \varepsilon a^2 + \frac{1}{4\varepsilon} b^2$)

$$\begin{aligned} (\nu - M \|\nabla \mathbf{v}_h\|) \|\nabla \phi_h\|^2 \leq\ & \varepsilon \|\nabla \phi_h\|^2 \\ & + \frac{\nu^2}{4\varepsilon} \|\nabla \eta\|^2 \\ & + \varepsilon \|\nabla \phi_h\|^2 \\ & + \frac{d}{4\varepsilon} \|p q_h\|^2 \\ & + \varepsilon \|\nabla \phi_h\|^2 \\ & + \frac{M^2}{4\varepsilon} (\|\mathbf{v}\| + \|\mathbf{v}_h\|)^2 \|\nabla \eta\|^2 . \end{aligned}$$

It follows

$$(\nu - M \|\nabla \mathbf{v}_h\| - 3\varepsilon) \|\nabla \phi_h\|^2 \leq \frac{\nu^2}{4\varepsilon} \|\nabla \eta\|^2 + \frac{d}{4\varepsilon} \|p q_h\|^2 + \varepsilon \|\nabla \phi_h\|^2 + \frac{M^2}{4\varepsilon} (\|\mathbf{v}\| + \|\mathbf{v}_h\|)^2 \|\nabla \eta\|^2 .$$

Small data condition and $\nu \|\nabla \mathbf{v}_h\| \leq \|f\|_{-1}$ yield

$$\begin{aligned} M \|\nabla \mathbf{v}_h\| &\leq MRe \|f\|_{-1} \\ &\leq \alpha \frac{1}{Re} \\ &\leq \frac{1}{Re}. \end{aligned}$$

Thus, we obtain, assuming $\varepsilon = \frac{1-\alpha}{6} \nu \|\nabla \phi_h\|$

$$\begin{aligned} \frac{\nu(1-\alpha)}{2} \|\nabla \phi_h\|^2 \leq\ & \frac{3\nu}{2(1-\alpha)} \|\nabla \eta\|^2 \\ & + \frac{3d}{2(1-\alpha)\nu} \|p - q_h\|^2 \\ & + \frac{3}{2(1-\alpha)\nu} \left(\frac{2}{\nu} \|f\|_* \right)^2 \|\nabla \eta\|^2 \end{aligned}$$

and further

$$\begin{aligned}\|\nabla\phi_h\|^2 \leq{}& \frac{3}{(1-\alpha)^2}\|\nabla\eta\|^2 \\ &+ \frac{3d}{(1-\alpha)^2\nu^2}\|p-q_h\|^2 \\ &+ \frac{12}{(1-\alpha)^2\nu^4}\|f\|_{-1}^2\|\nabla\eta\|^2 .\end{aligned}$$

Step 5: So we obtain

$$\|\nabla(\mathbf{v}-\mathbf{v}_h)\| \leq \|\nabla\eta\| + \|\nabla\phi_h\| .$$

The result follows with Theorem 3.16. □

Theorem 3.16 (Convergence of the FEM) *Suppose the global uniqueness condition*

$$\frac{M}{\nu^2}\|f\|_{-1} \leq \alpha < 1,$$

then it holds

$$\|\nabla(\mathbf{v}-\mathbf{v}_h)\| \leq C(\nu,f)\left\{\inf_{\varphi_h\in V_h}\|\nabla(\mathbf{v}-\varphi_h)\| + \inf_{q_h\in Q_h}\|p-q_h\|\right\}.$$

4 Time dependent Navier-Stokes equations

Consider the flow of a fluid in a region $\Omega \subset \mathbb{R}^2$ or $\Omega \subset \mathbb{R}^3$ bounded by walls and driven by a body force $f(x,t)$. The fluid velocity and pressure are functions $\mathbf{v}(x,t)$, $p(x,t)$ for $x \in \bar{\Omega}$, $0 \le t \le T$, which satisfy

$$\begin{cases} \partial_t \mathbf{v} + (\mathbf{v}\cdot\nabla)\mathbf{v} - \nu\Delta\mathbf{v} + \nabla p = f, & x \in \Omega,\ 0 < t \le T \\ \text{div } \mathbf{v} = 0, & x \in \Omega,\ 0 < t \le T \\ \mathbf{v}(x,0) = \mathbf{v}_0(x), & x \in \Omega \\ \mathbf{v} = 0, & \text{on } \partial\Omega \\ \int_\Omega p \mathrm{d}x = 0, & 0 < t \le T \end{cases} . \tag{4.1}$$

The key idea in making progress in the mathematical understanding of the NSE is the notion of weak solutions. If $(\mathbf{v}, p)$ is a *smooth* solution to (4.1), then multiplying the momentum equation by $\mathbf{v}$, integrating over Ω, integrating by parts and integrating in time gives

$$\underbrace{\frac{1}{2}\|\mathbf{v}(t)\|^2}_{\text{kinetic energy}} + \underbrace{\int_0^t \nu \left\|\nabla \mathbf{v}(t')\right\|^2 \mathrm{d}t'}_{\text{total energy dissipated over } (0,t)} = \underbrace{\frac{1}{2}\|\mathbf{v}_0\|^2}_{\text{initial kinetic energy}} + \underbrace{\int_0^t \left(f(t'), \mathbf{v}(t')\right) \mathrm{d}t'}_{\text{total power input}} .$$

Definition 4.1 (Function spaces) *Consider*

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

Then

$$L^2\left(0,T; H_0^1(\Omega)\right) := \left\{ v(x,t) : [0,T] \to H_0^1(\Omega) : \int_0^T \|\nabla v\|^2 \mathrm{d}t < \infty \right\},$$

$$L^2\left(0,T; L^2(\Omega)\right) := \left\{ v(x,t) : [0,T] \to L^2(\Omega) : \sup_{0<t\le T} \|\nabla v\|^2 < \infty \right\},$$

$$L^2\left(0,T; L_0^2(\Omega)\right) := \left\{ q(x,t) : [0,T] \to L_0^2(\Omega) : \int_0^T \|q\|^2 \mathrm{d}t < \infty \right\}.$$

Definition 4.2 (Strong solution of the Navier-Stokes equations) $(\mathbf{v}, p)$ *is a* strong solution *of* (4.1)*, if*

1. $\mathbf{v} \in L^2(0,T;X) \cap L^\infty\left(0,T; L^2(\Omega)\right)$ *with*

$$X = \left\{ \mathbf{v} \in L^2(\Omega)^d : \quad \nabla\mathbf{v} \in L^2(\Omega)^{d\times d},\ \mathbf{v} = 0 \text{ on } \partial\Omega \right\}.$$

2. $\mathbf{v} : [0,T] \to X$ *is a differentiable map and* $p : [0,T] \to Q$ *is an integrable map.*

3. *For every* $t' \in (0,T]$, $(\mathbf{v},p)$ *satisfies*

$$\int_0^{t'} (\partial_t \mathbf{v}, \varphi) + (\mathbf{v}\cdot\nabla\mathbf{v}, \varphi) - (p, \operatorname{div}\,\varphi) + \nu\,(\nabla\mathbf{v}, \nabla\varphi)\,\mathrm{d}t = \int_0^{t'} (f,\varphi)\,\mathrm{d}t$$

for all $\varphi \in L^2(0,T;X) \cap L^\infty\left(0,T;L^2(\Omega)\right)$ *and*

$$\int_0^{t'} (q, \operatorname{div}\,\mathbf{v})\,\mathrm{d}t = 0$$

for all $q \in L^2\left(0,T;L_0^2(\Omega)\right)$.

4. $\mathbf{v}_0 \in X \cap \{(q, \operatorname{div}\,\mathbf{v}_0) = 0\ \forall q \in Q\}$.

5. $\mathbf{v} \in L^4(0,T;X)$.

Theorem 4.3 (Uniqueness of strong solution) *Let* $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$. *The Navier-Stokes equations have* at most one *strong solution* $\mathbf{v}$ *(with* $\nabla\mathbf{v} \in L^4(0,T;L^2(\Omega))$*).*

Remark 4.4 (Clay prize) The one who proves (or proposes a counter-example to) the *existence* of the strong solution to (4.1), wins the *Clay prize* which is endowed with $1,000,000!!

The notion of weak solutions is due to Leray who called them *turbulent solutions*.

$$(\mathbf{v}(T),\phi(T)) + \int_0^T \{-(\mathbf{v},\partial_t\phi) + \nu\,(\nabla\mathbf{v},\nabla\phi) + (\mathbf{v}\cdot\nabla\mathbf{v},\phi)\}\,\mathrm{d}t = \int_0^T (f,\phi)\,\mathrm{d}t + (\mathbf{v}(0),\phi(0)) \tag{4.2}$$

for any function $\phi(x,t)$ which is smooth enough, vanishes on $\partial\Omega$ and satisfies

$$\operatorname{div}\,\phi = 0.$$

To give the precise definition of a weak solution we require the introduction of new mathematical structures.

Definition 4.5 (Functions with compact support) *The* support *of* ϕ *in* Ω, *supp*(ϕ), *is*

$$\mathit{supp}(\phi) := \overline{\{x \in \Omega : \quad \phi(x) \neq 0\}}.$$

A function ϕ *has* compact support *in* Ω, *if supp*(ϕ) *is a closed and bounded (i.e. compact) subset of* Ω.

$$\begin{aligned} D(\Omega) &:= \left\{\psi \in C^\infty(\Omega)^d : \quad \psi \text{ has compact support in } \Omega,\ \operatorname{div}\,\psi = 0 \text{ in } \Omega\right\},\\ H(\Omega) &:= \text{ completion of } D(\Omega) \text{ in } L^2(\Omega)^d,\\ D_T &:= \{\phi(x,t) \in C^\infty(\Omega\times[0,T]) : \quad \phi(x,t) \in D(\Omega)\ \forall 0 \le t \le T\}. \end{aligned}$$

Lemma 4.6 *The space* $H(\Omega)$ *can be characterized by*

$$H(\Omega) = \left\{\mathbf{v} \in L^2(\Omega)^d : \quad \operatorname{div}\,\mathbf{v} = 0 \text{ and } \mathbf{v}\cdot\mathbf{n} = 0 \text{ on } \partial\Omega\right\}.$$

Proof See, e.g., book by Galdi: Navier-Stokes, or Sohr: Navier-Stokes. □

Definition 4.7

$$L^2(0,T;V) := \left\{ \mathbf{v}(t) : [0,T] \to V : \quad \int_0^T \|\nabla \mathbf{v}\|^2 \, \mathrm{d}t < \infty \right\}.$$

Definition 4.8 (Weak solution) *Let* $\mathbf{v}_0 \in H(\Omega)$, $f \in L^2(\Omega \times (0,T))$. *A function*

$$\mathbf{v}(x,t) : \Omega \times [0,T] \to \mathbb{R}^d$$

is a weak solution *of* (4.1) *if*

1. $\mathbf{v} \in L^2(0,T;V) \cap L^\infty(0,T;H(\Omega))$.
2. $\mathbf{v}$ *satisfies the integral relation*

$$(\mathbf{v} \cdot \nabla \mathbf{u}, \mathbf{w}) \leq C \sqrt{\|\mathbf{v}\| \, \|\nabla \mathbf{v}\|} \, \|\nabla \mathbf{u}\| \, \|\nabla \mathbf{w}\|$$

   and (4.2) *for all* $\phi \in D_T$.

3. $$\frac{1}{2} \|\mathbf{v}(t)\|^2 + \nu \int_0^t \left\|\nabla \mathbf{v}(t')\right\|^2 \mathrm{d}t' \leq \frac{1}{2} \|\mathbf{v}_0\|^2 + \int_0^t \big(\mathbf{v}(t'), f(t')\big) \, \mathrm{d}t'.$$

In 2D it can be shown that a weak solution *exists* and is *unique*. In 3D there are slightly more restrictive concepts which are known to be unique but for which existence is unknown.

Conjecture [Leray] The lack of a uniqueness proof for weak solutions in 3D is not due to a weakness of mathematical techniques but is rather an *essential feature*. □

5 Saddle-point problem

5.1 General setting

Let X, M be Hilbert spaces, $a : X \times X \to \mathbb{R}$ a bilinear form and $b : X \times M \to \mathbb{R}$ another bilinear form.

Problem 5.1 (Saddle-point problem) *Find $u \in X$ such that*

$$\begin{aligned} J(u) &= \min_{v \in X} J(v), \\ b(u, \mu) &= \langle g, \mu \rangle \quad \forall \mu \in M, \end{aligned}$$

whereas

$$J(v) := \frac{1}{2} a(v, v) - \langle f, v \rangle .$$

The corresponding *Lagrange function* to the above *constrained optimization problem* reads:

$$\mathcal{L}(u, \lambda) := J(u) + \{ b(u, \lambda) - \langle g, \lambda \rangle \} .$$

The *necessary conditions* for an optimal solution are obviously

$$\begin{cases} \partial_u \mathcal{L}(u, \lambda) & = 0 \\ \partial_\lambda \mathcal{L}(u, \lambda) & = 0 \end{cases} .$$

These conditions are possibly not sufficient!

Problem 5.2 *Find $(u, \lambda) \in X \times M$ such that*

$$\begin{cases} a(u, v) + b(v, \lambda) & = \langle f, v \rangle \quad \forall v \in X \\ b(u, \mu) & = \langle g, \mu \rangle \quad \forall \mu \in M \end{cases} .$$

Example 5.3

$$\min_{x, y \in \mathbb{R}} x^2 + y^2 \quad s.t. \quad x + y = 2.$$

The Lagrange function reads:

$$\mathcal{L}(x, y, \lambda) = x^2 + y^2 + \lambda(x + y - 2).$$

From the necessary optimality conditions it follows that

$$\begin{aligned} x = y &= 1, \\ \lambda &= -2, \end{aligned}$$

is a (*the*) saddle point.

Example 5.4

$$\min_{x,y\in\mathbb{R}} x^2 + y^2$$

s.t.

$$\begin{cases} x + y & = 2 \\ 3x + 3y & = 6 \end{cases}$$

Lagrange function:

$$\mathcal{L}(x, y, \lambda, \mu) = x^2 + y^2 + \lambda(x + y - 2) + \mu(3x + 3y - 6).$$

This problem is *not well-posed*! First, the solution is not unique since we have the general relation

$$\lambda + 3\mu = -2.$$

Second, if we disturb the second constraint

$$3x + 3y = 6.000000000000001,$$

there exists no solution because the constraints are contradicting! So the solution is not stable under small perturbations.

5.1.1 Discrete setting

Let X_h, M_h be Hilbert spaces, $a : X_h \times X_h \to \mathbb{R}$ a bilinear form and $b : X_h \times M_h \to \mathbb{R}$ another bilinear form.

Problem 5.5 (Saddle-point problem) *Find $u_h \in X_h$ such that*

$$\begin{aligned} J(u_h) &= \min_{v_h \in X_h} J(v_h), \\ b(u_h, \mu_h) &= \langle g, \mu_h \rangle \quad \forall \mu_h \in M_h, \end{aligned}$$

whereas

$$J(v) := \frac{1}{2} a(v, v) - \langle f, v \rangle .$$

The corresponding *Lagrange function* to the above *discrete constrained optimization problem* reads:

$$\mathcal{L}(u_h, \lambda_h) := J(u_h) + \{b(u_h, \lambda_h) - \langle g, \lambda_h \rangle\} .$$

The *necessary conditions* for an optimal solution are obviously

$$\begin{cases} \partial_{u_h} \mathcal{L}(u_h, \lambda_h) & = 0 \\ \partial_{\lambda_h} \mathcal{L}(u_h, \lambda_h) & = 0 \end{cases} .$$

These conditions are possibly not sufficient!

Problem 5.6 *Find $(u_h, \lambda_h) \in X_h \times M_h$ such that*

$$\begin{cases} a(u_h, v_h) + b(v_h, \lambda_h) & = \langle f, v_h \rangle \quad \forall v_h \in X_h \\ b(u_h, \mu_h) & = \langle g, \mu_h \rangle \quad \forall \mu_h \in M_h \end{cases} .$$

5.2 The inf − sup condition

Let $A : X \to X'$ be an operator where X' denotes the dual space of X.

$$\langle Au, v\rangle = a(u, v) \quad \forall v \in X$$

Different notation:

$$\langle Au, v\rangle = (Au)(v) = a(u, v).$$

Let $B : X \to M'$ be the operator defined by

$$\langle Bu, \mu\rangle = (Bu)(\mu) = b(u, \mu) \quad \forall \mu \in M$$

and $B' : M \to X'$ the adjoint operator

$$\langle B'\lambda, v\rangle = (B'\lambda)(v) = b(v, \lambda) \quad \forall v \in X.$$

With these notations, Problem 5.2 is equivalent to

$$\begin{cases} Au + B'\lambda &= f \\ Bu &= g \end{cases} \tag{5.1}$$

For finite dimensional spaces X_h, M_h a representation of (5.1) can be obtained by means of a matrix formulation, i.e.,

$$\begin{pmatrix} A_h & B_h^\top \\ B_h & 0 \end{pmatrix} \begin{pmatrix} u_h \\ \lambda_h \end{pmatrix} = \begin{pmatrix} f_h \\ g_h \end{pmatrix} \tag{5.2}$$

whereas $A_h \in \mathbb{R}^{n\times n}$ and $B_h \in \mathbb{R}^{n\times m}$.

5.2.1 The Stokes equations

Let $X = \left(H_0^1(\Omega)\right)^n$, $M = L_0^2(\Omega)$ and

$$a(u, v) := \int_\Omega \nabla u \nabla v \mathrm{d}x,$$
$$b(v, q) := -\int_\Omega q \mathrm{div}\ v \mathrm{d}x.$$

Then the Stokes problem reads:

Problem 5.7 (Stokes) *Find $(u, p) \in X \times M$ such that*

$$\begin{cases} a(u, v) + b(v, q) &= (f, v)_0 \quad \forall v \in X \\ b(u, q) &= 0 \quad \forall q \in M. \end{cases}$$

5.2.2 Laplace equation

$$-\Delta u = f \tag{5.3}$$

can be reformulated to

$$\begin{cases} \nabla u &= \sigma \\ \mathrm{div}\ \sigma &= -f \end{cases}$$

since

$$\Delta u = \mathrm{div}\ \underbrace{(\nabla u)}_{=:\sigma} = -f.$$

The corresponding variational formulation reads as follows:

Problem 5.8 *Find* $(\sigma, u) \in L^2(\Omega)^d \times H_0^1(\Omega)$

$$\begin{cases} (\sigma,\tau)_0 - (\tau,\nabla u)_0 &= 0 \quad \forall \tau \in L^2(\Omega)^d \\ -(\sigma,\nabla v)_0 &= -(f,v)_0 \quad \forall v \in H_0^1(\Omega) \end{cases}.$$

With $X := L^2(\Omega)^d$, $M := H_0^1(\Omega)$ and

$$\begin{aligned} a(\sigma,\tau) &:= (\sigma,\tau)_0\,, \\ b(\tau,v) &:= -(\tau,\nabla v)_0\,, \end{aligned}$$

Laplace's equation (5.3) is in fact a problem of the form of Problem 5.2 and, therefore, a saddle-point problem.

Remark 5.9 1. σ is the *tension* in the physical system.

2. (5.3) is also equivalent to
$$\min_{u\in H_0^1(\Omega)} \frac{1}{2}(\nabla u,\nabla u) - \langle f,u\rangle$$
which is an *unconstrained* optimization problem.

The optimization problem related to Problem 5.8 is defined by

$$\begin{cases} \min_{\sigma\in L^2(\Omega)^d} \frac{1}{2}(\sigma,\sigma) & \\ \text{subject to} & -(\sigma,\nabla v)_0 = -(f,v)_0 \quad \forall v \in H_0^1(\Omega) \end{cases}.$$

This is a so called *mixed formulation*.

5.2.3 The $\inf-\sup$ condition

Theorem 5.10 *The following statements are all equivalent:*

1. $\exists \beta > 0$ *with*
$$\inf_{\mu\in M}\sup_{v\in X}\frac{b(\mu,\nu)}{\|v\|\,\|\mu\|} \geq \beta. \tag{5.4}$$

2. *The operator* $B : V^\perp \to M'$ *is an isomorphism and*
$$\|Bv\| \geq \beta\,\|v\| \quad \forall v \in V^\perp,$$
whereas
$$V := \{v \in X : \quad b(v,\mu) = 0\ \forall \mu \in M\}\,.$$

3. *The operator* $B' : M \to V^0$ *is an isomorphism and*
$$\left\|B'\mu\right\| \geq \beta\,\|\mu\| \quad \forall \mu \in M.$$

Definition 5.11 *A finite element fulfils the so called Babuška-Brezzi (Nečas) condition, if the following properties hold:*

1. *The bilinear form* a *is* V*-elliptic.*
2. *The* $\inf-\sup$ *condition* (5.4) *holds.*

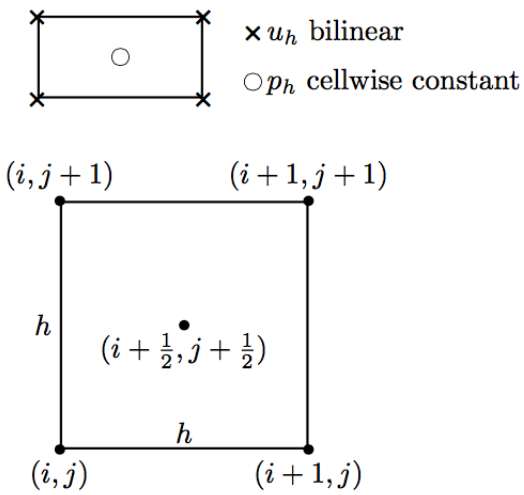

Figure 5.1: Unstable $Q_1 - P_0$ element

Remark 5.12 We have already seen the inf – sup condition in (3.7) or Definition 3.6.

Example 5.13 (Unstable elements) We are considering the so called $Q_1 - P_0$ element on a rectangle, i.e., for the velocity we consider bilinear functions with degree of freedoms at the vertices of the rectangle and for the pressure we consider constant functions whose function value we define at the center of the rectangle (cf. Figure 5.1):

$$\begin{aligned} X_h &:= \left\{ v \in C^0(\bar{\Omega})^2 : \quad v_{|T} \in Q_1 \text{ also bilinear for } T \in \mathcal{T}_h \right\}, \\ M_h &:= \left\{ q \in L_0^2(\Omega) : \quad q_{|T} \in P_0 \text{ for } T \in \mathcal{T}_h \right\}. \end{aligned}$$

With these definitions we can compute

$$\begin{aligned} \int_{T_{ij}} q \operatorname{div} v \mathrm{d}x &= h^2 q_{i+\frac{1}{2},j+\frac{1}{2}} \operatorname{div} v_{i+\frac{1}{2},j+\frac{1}{2}} \\ &= h^2 q_{i+\frac{1}{2},j+\frac{1}{2}} \frac{1}{2h} \left[u_{i+1,j+1} + u_{i+1,j} - u_{i,j+1} - u_{i,j} + w_{i+1,j+1} + w_{i,j+1} - w_{i+1,j} - w_{i,j} \right] \end{aligned}$$

assuming $v = (u, w)^\top$.

$$\begin{aligned} \int_\Omega q \operatorname{div} v \mathrm{d}x &= h^2 \sum_{i,j} \left[u_{i,j} (\nabla_1 q)_{i,j} + w_{i,j} (\nabla_2 q)_{i,j} \right], \\ (\nabla_1 q)_{i,j} &= \frac{1}{2h} \left[q_{i+\frac{1}{2},j+\frac{1}{2}} + q_{i+\frac{1}{2},j-\frac{1}{2}} - q_{i-\frac{1}{2},j+\frac{1}{2}} - q_{i-\frac{1}{2},j-\frac{1}{2}} \right], \\ (\nabla_2 q)_{i,j} &= \frac{1}{2h} \left[q_{i+\frac{1}{2},j+\frac{1}{2}} + q_{i-\frac{1}{2},j+\frac{1}{2}} - q_{i+\frac{1}{2},j-\frac{1}{2}} - q_{i-\frac{1}{2},j-\frac{1}{2}} \right]. \end{aligned}$$

We have $q \in \ker(B_h')$ if

$$\int_\Omega q \operatorname{div} v \mathrm{d}x = 0 \quad \forall v \in X_h,$$

i.e., if and only if $\nabla_1 q$ and $\nabla_2 q$ are equal to zero for inner nodes of the triangulation:

$$q_{i+\frac{1}{2},j+\frac{1}{2}} = q_{i-\frac{1}{2},j-\frac{1}{2}}, \quad q_{i+\frac{1}{2},j-\frac{1}{2}} = q_{i-\frac{1}{2},j+\frac{1}{2}}.$$

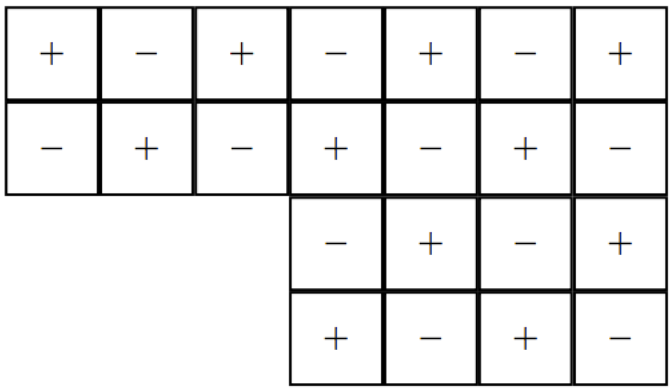

Figure 5.2: Pressure instability in L-shaped domain

A possible setup for these equations is to assume

$$q_{i+\frac{1}{2},j+\frac{1}{2}} = \begin{cases} a & \text{for } i+j = 2k \\ b & \text{for } i+j = 2k+1 \end{cases}, k \in \mathbb{N},$$

cf. Figure 5.2.

Theorem 5.14 (Fortin Criterion) *Let $b : X \times M \to \mathbb{R}$ a bilinear form which fulfils the* $\inf - \sup$ *condition. Let $X_h \subset X$ and $M_h \subset M$ be finite-dimensional subspaces. Let $\Pi_h : X \to X_h$ be a linear and bounded operator, such that*

$$b(v - \Pi_h v, \mu_h) = 0 \quad \forall \mu_h \in M_h.$$

If $\|\Pi_h\| \leq C$ where C is independent from h, then (X_h, M_h) is a stable finite element space.

Proof

$$\begin{aligned} \beta \|\mu_h\| &\leq \sup_{v \in X} \frac{b(v, \mu_h)}{\|v\|} \\ &= \sup_{v \in X} \frac{b(\Pi_h v, \mu_h)}{\|v\|} \\ &\leq C \sup_{v \in X} \frac{b(\Pi_h v, \mu_h)}{\|\Pi_h v\|} \\ &= C \frac{b(v_h, \mu_h)}{\|v_h\|}. \end{aligned}$$

□

Example 5.15 (MINI element) The space for the velocity is defined as

$$\mathcal{M}^1_{0,0} := \left\{ v_h \in C(\bar{\Omega}) \cap H^1_0(\Omega)^d : \quad v_{h|T} \in \mathcal{P}_1 \text{ for } T \in \mathcal{T}_{\frac{h}{2}} \right\}$$

plus a *bubble function*, e.g.,

$$b(x) = \lambda_1 \lambda_2 \lambda_3$$

in barycentric representation, i.e.

$$X_h := \mathcal{M}^1_{0,0} \bigoplus B_3,$$

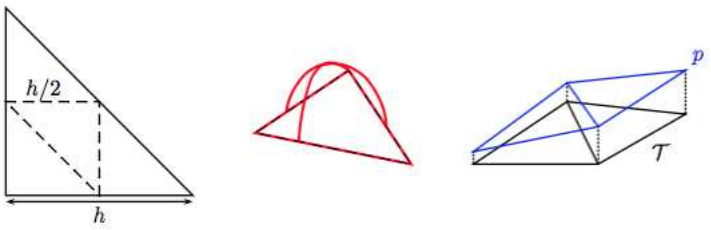

Figure 5.3: MINI Element

where B_3 corresponds to the space spanned by the bubble function

$$B_3 := \left\{ v \in C^0(\bar{\Omega}) : \quad v_{|T} \in span\{\lambda_1\lambda_2\lambda_3\} \text{ for } T \in \mathcal{T}_h \right\}.$$

The space for the pressure is defined as

$$M_h := \left\{ q_h \in C(\Omega) \cap L_0^2(\Omega) : \quad q_{h|T} \in \mathcal{P}_1 \text{ for } T \in \mathcal{T}_h \right\}.$$

Lemma 5.16 (MINI element) *Let Ω be convex and $\partial\Omega$ smooth enough. Then the MINI element (X_h, M_h) is stable.*

Proof We define the projector

$$\Pi_h^0 : H_0^1(\Omega) \to \mathcal{M}_{0,0}^1$$

by means of the solution of the Helmholtz equation:

$$\left(\nabla\left(\Pi_h^0 v\right), \nabla w\right)_0 + \left(\Pi_h^0 v, q\right)_0 = (\nabla v, w)_0 + (v, w)_0 \quad \forall w \in \mathcal{M}_{0,0}^1.$$

Obviously,

$$\begin{aligned} \left\|\Pi_h^0 v\right\|_1 &\leq \|v\|_1, \\ \left\|\Pi_h^0 v - v\right\|_0 &\leq C_1 h \left\|\Pi_h^0 v - v\right\|_1 \leq C_2 h \|v\|_1. \end{aligned}$$

Further, we define $\Pi_h^1 : L^2(\Omega) \to B_3$ by

$$\int_T \left(\Pi_h^1 v - v\right) \mathrm{d}x = 0 \quad \forall T \in \mathcal{T}_h.$$

One can easily show that $\left\|\Pi_h^1 v\right\|_0 \leq C_3 \|v\|_0$. Let

$$\Pi_h v := \Pi_h^0 v + \Pi_h^1 \left(v - \Pi_h^0 v\right).$$

Obviously,

$$\int_T \left(\Pi_h v - v\right) \mathrm{d}x = 0 \quad \forall T \in \mathcal{T}_h.$$

Further,

$$\begin{aligned} b(v - \Pi_h v, q_h) &= \int_\Omega \text{div } (v - \Pi_h v) q_h \mathrm{d}x \\ &= \int_{\partial\Omega} (v - \Pi_h v) \mathbf{n} q_h \mathrm{d}S - \int_\Omega (v - \Pi_h v) \nabla q_h \mathrm{d}x \\ &= 0 \quad \forall q_h \in M_h. \end{aligned}$$

Additionally,

$$\begin{aligned}\|\Pi_h v\|_1 &\leq \left\|\Pi_h^0 v\right\|_1 + C_4 h^{-1} \left\|\Pi_h^1 (v - \Pi_h^0 v)\right\|_0 \\ &\leq C_2 \|v\|_1 + C_4 h^{-1} C_3 \left\|v - \Pi_h^0 v\right\|_0 \\ &\leq C_2 \|v\|_1 + C_4 C_3 C_2 \|v\|_1 .\end{aligned}$$

□

6 A posteriori error estimation

6.1 Introduction

Find $u \in H_0^1(\Omega)$ such that

$$(\nabla u, \nabla \varphi) = (f, \varphi) \quad \forall \varphi \in H_0^1(\Omega).$$

The corresponding discrete problem reads: Find $u_h \in V_h \subset H_0^1(\Omega)$ such that

$$(\nabla u_h, \nabla \varphi) = (f, \varphi) \quad \forall \varphi \in V_h.$$

We want to know or at least estimate the error

$$\|u - u_h\| .$$

Of course, the concrete error depends on the considered norm. We know, for example, that

$$\|u - u_h\|_1 \leq \frac{C}{\alpha} \inf_{v \in V_h} \|u - v\|_1 \leq \frac{C}{\alpha} \|u - I_h u\|_1 \leq C(u, u_h)$$

due to Céa's lemma (and interpolation error estimation as well as the assumption of C^2 regularity). *Is this an a priori or an a posteriori error estimation?* This is an *a priori* error estimation because we estimate the error with the *unknown* function u. An estimation of the form

$$\|u - u_h\| \leq C(u_h)$$

is called an *a posteriori* error estimation because we estimate the error with the *known* discrete solution.

6.2 Residual based error estimator

We consider the Laplace equation in the following form:

$$\begin{cases} -\Delta u &= f \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega \end{cases}.$$

The variational formulation reads: Find $u \in H_0^1(\Omega)$ or $u_h \in V_h \subset H_0^1(\Omega)$ such that

$$\begin{aligned} (\nabla u, \nabla \varphi) &= (f, \varphi) && \forall \varphi \in H_0^1(\Omega), \\ (\nabla u_h, \nabla \varphi) &= (f, \varphi) && \forall \varphi \in V_h. \end{aligned}$$

We now define a *goal functional* $J(u)$ (e.g., $J(u) := u(x_{NumPDE3})$ if we want to know the function value at a specific point) and define the error as

$$e_h := J(u) - J(u_h).$$

In order to compute the error with respect to this goal functional, we consider an additional problem called *dual problem*: Find $z \in H_0^1(\Omega)$ such that

$$(\nabla z, \nabla \varphi) = (J(u), \varphi) \quad \forall \varphi \in H_0^1(\Omega).$$

The discrete dual problem then reads: Find $z_h \in V_h$ such that

$$(\nabla z_h, \nabla \varphi) = (J(u), \varphi) \quad \forall \varphi \in V_h.$$

The function z is called *weight function*. In the case of Poisson's equation it is in fact *Green's function*. The weight function is a measure for the *sensitivity* of the solution u of the *primal problem* with respect to the goal functional.

Lemma 6.1

$$J(u) - J(u_h) = (\nabla u, \nabla (z - I_h z)) + h.o.t.$$

where h.o.t. denotes higher order terms of the order $\mathcal{O}\left(h^k\right)$, $k \geq 2$.

$$\begin{aligned}
J(u) - J(u_h) &\leq (\nabla u_h, \nabla (z - I_h z)) \\
&\leq \int_\Omega \nabla u_h \cdot \nabla (z - I_h z) \, \mathrm{d}x \\
&\leq \sum_{T \in \mathcal{T}_h} \left| \int_T \nabla u_h \cdot \nabla (z - I_h z) \, \mathrm{d}x \right| \\
&\leq \sum_{T \in \mathcal{T}_h} \left\{ \left| \int_T (-\Delta u_h, z - I_h z) \, \mathrm{d}x \right| + |\lfloor \nabla u_h \rfloor_{\partial T}| \, |\lfloor z - I_h z \rfloor_{\partial T}| \right\}
\end{aligned}$$

This estimation enables us to develop a *refinement strategy*: Cells T with "large" contribution to the error estimation are refined, whereas cells with "small" contribution are kept.

We define the following measures in order to estimate the *local* error in the discrete solution:

$$\begin{aligned}
R_T &:= R_T(u_h) := \Delta u_h + f \quad \text{for} \quad T \in \mathcal{T}_h \text{ (cell error)}, \\
R_e &:= R_e(u_h) := \lceil \partial_{\mathbf{n}} u_h \rceil_e \quad \text{for} \quad e \in \Gamma_h \text{ (jump error)},
\end{aligned}$$

where

$$\begin{aligned}
\lceil \partial_{\mathbf{n}} u_h \rceil_e &:= - \left(\nabla u_{h,T} \mathbf{n}_T + \nabla u_{h,T'} \mathbf{n}_{T'} \right) \\
&= - \left(\nabla u_{h,T} \mathbf{n}_T - \nabla u_{h,T'} \mathbf{n}_T \right) \\
&= \left(\nabla u_{h,T'} - \nabla u_{h,T} \right) \mathbf{n}_T
\end{aligned}$$

and T' denotes the neighbouring cell of T adjacent to edge e. R_e is a measure for the discontinuity of the gradient of the discrete solution between two neighbouring cells.

Excurse *Goal:* Solve the linear system $Ax = b$, $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$. Assume that x_k is an approximation of x.

$$A(x - x_k) = b - Ax_k.$$

Let $r_k = b - Ax_k$ be the residual and $e_k = x - x_k$ the error. Then the above equation is equivalent to

$$\begin{aligned}
&\Leftrightarrow && Ae_k = r_k \\
&\Leftrightarrow && e_k = A^{-1} r_k \quad \text{(assuming regularity of } A\text{)} \\
&\Rightarrow && \|e_k\| \leq \left\|A^{-1}\right\| \|r_k\| .
\end{aligned}$$

PICTURES!!!

Figure 6.1: Error contributions

The goal of a posteriori error estimators is to give a *computable* approximation of $\left\|A^{-1}\right\|$. □

$$\begin{aligned}\omega_T &:= \left\{T' \in \mathcal{T}_h : \quad T \text{ and } T' \text{ have a common edge or } T = T'\right\},\\ \omega_e &:= \left\{T' \in \mathcal{T}_h : \quad e \subset \partial T'\right\}.\end{aligned}$$

6.2.1 Residual estimators

These estimators were first developed by Babuška & Rheinhold [1978].

We consider the local expression

$$\eta_{T,R} := \left\{ h_T^2 \left\|R_T\right\|_{0,T}^2 + \frac{1}{2}\sum_{e\in\partial T} h_e \left\|R_e\right\|_{0,e}^2 \right\}^{\frac{1}{2}} \quad \text{for} \quad T \in \mathcal{T}_h.$$

Based on this local description of the error, we define the global estimator

$$\eta_R := \left\{ \sum_{T\in\mathcal{T}_h} \left\|R_T\right\|_{0,T}^2 + \sum_{e\in\Gamma_h} h_e \left\|R_e\right\|_{0,e}^2 \right\}^{\frac{1}{2}}.$$

Theorem 6.2 *Let $\mathcal{T}_h$ be a quasiuniform triangulation. Then a constant $C = C(\Omega,\kappa)$ exists such that*

$$\left\|u - u_h\right\|_{1,\Omega} \le C \left(\sum_{T\in\mathcal{T}_h} \eta_{T,R}^2 \right)^{\frac{1}{2}}$$

and

$$\eta_{T,R} \le C \left\{ \left\|u - u_h\right\|_{1,\omega_T}^2 + \sum_{T'\in\omega_T} h_{T'}^2 \left\|f - f_h\right\|_{0,T'}^2 \right\}^{\frac{1}{2}}$$

for all $T \in \mathcal{T}_h$.

Proof Duality argument:

$$\left|u - u_h\right|_1 \le \sup_{w\in H_0^1(\Omega),\left|w\right|_1=1} (\nabla(u - u_h), \nabla w)_0$$

and

$$(\nabla(u - u_h), \nabla v_h) = 0 \quad \forall v_h \in S_h,$$

i.e., Galerkin orthogonality.

We define the functional

$$l : H_0^1(\Omega) \to \mathbb{R}$$

by means of

$$\begin{aligned}\langle l, w\rangle &:= (\nabla(u-u_h), \nabla w)_{0,\Omega} \\ &= (f,w)_{0,\Omega} - \sum_T (\nabla u_h, \nabla w)_{0,T} \\ &= (f,w)_{0,\Omega} - \sum_T \left\{ (-\Delta u_h, w)_{0,T} + \sum_{e\subset\partial T} (\nabla u_h \cdot \mathbf{n}, w)_{0,e} \right\} \\ &= \sum_T (\Delta u_h + f, w)_{0,T} + \sum_{e\in\Gamma_h} (\lceil \partial_{\mathbf{n}} u_h \rceil, w)_{0,e} \\ &= \sum_T (R_T, w)_{0,T} + \sum_{e\in\Gamma_h} (R_e, w)_{0,e}\,.\end{aligned}$$

Considering the Interpolation operator of Clément (see Numerik PDE) we may write

$$\begin{aligned}\|w - I_h w\|_{0,T} &\le C h_T \|\nabla w\|_{0,\omega_T} \quad \text{for} \quad T\in\mathcal{T}_h, \\ \|w - I_h w\|_{0,e} &\le C h_e^{\frac{1}{2}} \|\nabla w\|_{0,\omega_e} \quad \text{for} \quad e\in\gamma_h\end{aligned}$$

and

$$\begin{aligned}\langle l, w\rangle &= \langle l, w - I_h w\rangle \\ &\le \sum_T \|R_T\| \, \|w - I_h w\|_{0,T} + \sum_{e\in\Gamma_h} \|R_e\|_{0,e} \, \|w - I_h w\|_{0,e} \\ &\le C \sum_T h_T \, \|R_T\|_{0,T} \, |w|_{1,\omega_T} + C \sum_{e\in\Gamma_h} h_e^{\frac{1}{2}} \, \|R_e\|_{0,e} \, |w|_{1,\omega_e} \\ &\le C \sum_T \eta_{T,R} \, |w|_{1,T} \\ &\le C \eta_R \, |w|_{1,\Omega}\,.\end{aligned}$$

Now the use of the duality argument yields the inequalities. □

6.2.2 Dual weighted Residual method (DWR)

Assume a *linear* (goal) functional $J(\cdot) : V \to \mathbb{R}$. We are interested in controlling the error

$$J(e_h) = J(u) - J(u_h).$$

In this context, we consider the dual problem: Find $z \in V$ $(= H_0^1(\Omega))$ such that

$$(\nabla z, \nabla\varphi) = J(\varphi) \quad \forall\varphi \in V.$$

Consider the error

$$J(e_h) = (\nabla e_h, \nabla z) = (\nabla eh, \nabla\,(z - \psi_h)) \quad \forall\psi_h \in V_h.$$

It follows that

$$\begin{aligned}J(e_h) &= \sum_{T\in\mathcal{T}_h} \{ -(\Delta e_h, z - \psi h)_T + (\partial_{\mathbf{n}} e_h, z - \psi_h)_{\partial T} \} \\ &= \sum_{T\in\mathcal{T}_h} \left\{ \left(\underbrace{f + \Delta u_h}_{\text{Residual}}, z - \psi_h \right)_T - (\partial_{\mathbf{n}} u_h, z - \psi_h)_{\partial T} \right\} \\ &= \sum_{T\in\mathcal{T}_h} \left\{ (f + \Delta u_h, z - \psi_h)_T - \frac{1}{2} (\lceil \partial_{\mathbf{n}} u_h \rceil, z - \psi_h)_{\partial T} \right\}.\end{aligned}$$

PICTURE!!!

Figure 6.2: Superconvergence

We have

$$|J(e_h)| \leq \sum_{T\in\mathcal{T}_h} \left| (f + \Delta u_h, z - \psi_h)_T - \frac{1}{2} \left(\lceil \partial_{\mathbf{n}} u_h \rceil, z - \psi_h \right)_{\partial T} \right| .$$

Considering the Hölder inequality

$$\begin{aligned} |J(e_h)| &\leq \sum_{T\in\mathcal{T}_h} \left\{ \underbrace{\|f + \Delta u_h\|_T}_{=:\rho_T(u_h)} \underbrace{\|z - I_h z\|_T}_{=:\omega_T(z)} + \frac{1}{2} \|\lceil \partial_{\mathbf{n}} u_h \rceil\|_{\partial T\setminus\partial\Omega} \|z - I_h z\|_{\partial T} \right\} \\ &\leq \sum_{T\in\mathcal{T}_h} \left\{ \rho_T(u_h)\omega_T(z) + \rho_{\partial T}(u_h)\omega_{\partial T}(z) \right\} , \end{aligned}$$

where

$$\begin{aligned} \rho_{\partial T}(u_h) &:= \frac{1}{2} h_T^{-\frac{1}{2}} \|\lceil \partial_{\mathbf{n}} u_h \rceil\|_{\partial T\setminus\partial\Omega} , \\ \omega_{\partial T}(z) &:= h_T^{\frac{1}{2}} \|z - I_h z\|_{\partial T} . \end{aligned}$$

The exact solution z to the dual problem is unknown, so we need a possibility to compute the error in the dual solution. Some ideas might be:

1. Take an approximation: Find $z_h \in V_h$ such that

   $$(\nabla z_h, \nabla\varphi) = J(\varphi) \quad \forall\varphi \in V_h.$$

   This is a bad idea, because the interpolation of the approximation is identical with the approximation itself so the "'error"' in the dual solution will be always zero and therefore the error estimator will always give the value zero.

2. A priori error estimation for z:

   $$\|z - I_h z\|_T \leq Ch \left\|\nabla^2 z\right\|_T .$$

   In this case the second derivative of z is not known. It possible to assume that this second derivative is constant and take Ch as an "'estimation"' for the dual solution but in fact the information of the dual solution is completely lost. Also bad idea!

3. Take an approximation *in a larger space*: Find $z_{\frac{h}{2}} \in V_{\frac{h}{2}}$ such that

   $$\left(\nabla z_{\frac{h}{2}}, \nabla\varphi\right) = J(\varphi) \quad \forall\varphi \in V_{\frac{h}{2}}.$$

   This is possible but increases the computational costs extremely! To lower the computational costs the refinement $h \to \frac{h}{2}$ is often done only locally.

4. Trick for higher derivatives: Locally refinement of the grid leads to *superconvergence*: Convergence order increases from h^2 to $h^{\frac{5}{2}}$.

6.3 Exact grid optimization

Optimization problem 1 Let

- N: number of cells
- η: error estimator

$$\min N \quad \text{such that} \quad \eta \leq TOL$$

with some tolerance TOL.

Optimization problem 2 Let

- N: number of cells
- η: error estimator

$$\min \eta \quad \text{such that} \quad N \leq N_{max}$$

with some maximum number of cells N_{max}.

We can have the following view:

$$h_T^{-1}\rho_T(u_h) = h_T^{-1}\left\|f + \Delta u_h\right\|_T + \frac{1}{2}h_T^{-\frac{3}{2}}\left\|h_T^{-1}\left\lceil \partial_n u_h \right\rceil\right\|_{\partial T\setminus\partial\Omega} \to \phi_u(x_T)$$
$$h_T^{-3}\omega_T(z_h) = \max\left\{h_T^{-3}\left\|z - I_h z\right\|_T, h_T^{-\frac{5}{2}}\left\|z - I_h z\right\|_{\partial T}\right\} \to \psi_z(x_T)$$

Let $A(x) = \phi(x)\psi(x)$.

$$\begin{aligned}\sum_{T\in\mathcal{T}_h}\rho_T(u_h)\omega_T(z_h) &= \sum_{T\in\mathcal{T}_h} h_T^4\left\{h_T^{-4}\rho_T(u_h)\omega_T(z_h)\right\}\\ &\approx \int_\Omega h^2(x)A(x)\mathrm{d}x.\\ N &= \sum_{T\in\mathcal{T}_h} 1\\ &= \sum_{T\in\mathcal{T}_h} h_T^2 h_T^{-2}\\ &\approx \int_\Omega h^{-2}(x)\mathrm{d}x.\end{aligned}$$

So the optimization problem reads:

$$\min F(h) := \int_\Omega h^2(x)A(x)\mathrm{d}x \quad \text{such that} \quad N(h) := \int_\Omega h^{-2}(x)\mathrm{d}x = N_{max}.$$

The Lagrange functional is

$$L(h,\lambda) := F(h) + \lambda(N(h) - N_{max}).$$

It holds

$$\mathrm{d}_t L(h + t\varphi, \lambda + t\mu)_{|t=0} = 0.$$

7 Practical aspects

7.1 Solution of the discrete Stokes problems

We only consider finite element discretizations which fulfil (3.7). With the usual nodal basis functions of the spaces $X_h \subset X$ and $Q_h \subset Q$ the discrete system can be written in block matrix form as

$$\mathcal{A}\xi = \begin{pmatrix} A & B \\ -B^\top & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix} =: \beta, \tag{7.1}$$

where $\xi := (x, y)^\top$ denotes the vector of nodal values of the representations

$$v_h = \sum_{i=1}^{N_X} x_i \psi_h^i, \quad p_h = \sum_{i=1}^{N_Q} y_i \chi_h^i,$$

and the corresponding matrices and right hand sides

$$A := \left(\left(\nabla \psi_h^j, \nabla \psi_h^i \right) \right)_{i,j=1}^{N_X}, \quad B := - \left(\left(\chi_h^j, \operatorname{div} \psi_h^i \right) \right)_{i,j=1}^{N_X, N_Q}, \quad b := \left(\left(f, \psi_h^i \right) \right)_{i=1}^{N_X}.$$

What possibilities do we have to solve the system $\mathcal{A}\xi = \beta$?

1. Apply our known methods to solve linear systems directly to $\mathcal{A}$, e.g., LU decomposition, CG, GMRES or multigrid methods.

   a) *Advantages:*

      i. We can reuse our existing implementations of these methods.

      ii. $\mathcal{A}$ is known to be regular due to the assumption (3.7).

   b) *Disadvantages:*

      i. What is the condition number of $\mathcal{A}$? The condition number of A is of the order $\mathcal{O}\left(h^{-2}\right)$ but what about the block system with the contributions of B?

      ii. The memory consumption is higher compared to storing A and B separately. If B cannot be transposed with low computational costs, it is beneficial, to store additionally $-B^\top$. In every case, the "0" does not need to be stored in memory.

2. Are there methods which only need methods to efficiently invert A, i.e., solve linear systems with the system matrix A, and multiplications with B (and maybe $-B^\top$)? The answer to this question is *yes* and we will discuss several possibilities in the next section.

7.1.1 Schur complement method

Let

$$\mathcal{A}\xi = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}$$

be a linear system in block matrix form. Assume that A is regular. Then the first block can be solved for x and the result inserted into the second block which yields

$$\begin{aligned} x &= A^{-1}f - A^{-1}By \\ \left(D - CA^{-1}B\right) y &= g - CA^{-1}f. \end{aligned}$$

The matrix $\Sigma := D - CA^{-1}B$ is called the *Schur complement* of A in the block matrix $\mathcal{A}$. With this notion of Σ, $\mathcal{A}$ can be decomposed in block triangular matrices:

$$\mathcal{A} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & 0 \\ C & D - CA^{-1}B \end{pmatrix} \cdot \begin{pmatrix} I & A^{-1}B \\ 0 & I \end{pmatrix}.$$

In case of the Stokes system, the above results reduce to

$$\begin{aligned} x &= A^{-1}b - A^{-1}By, \\ B^\top A^{-1}By &= B^\top A^{-1}b, \\ \Sigma &= B^\top A^{-1}B, \\ \mathcal{A} &= \begin{pmatrix} A & B \\ -B^\top & 0 \end{pmatrix} = \begin{pmatrix} A & 0 \\ -B^\top & \Sigma \end{pmatrix} \cdot \begin{pmatrix} I & A^{-1}B \\ 0 & I \end{pmatrix}. \end{aligned}$$

7.1.1.1 Uzawa algorithm

The "classical" method to solve the saddle-point problem (7.1) is the *Uzawa algorithm*:

Algorithm 7.1 (Uzawa) Given an initial value y^0. For $l \in \mathbb{N}$ compute:

1. $$Ax^l = b - By^{l-1}$$

2. $$y^l = y^{l-1} + \theta B^\top x^l$$

Until convergence is reached.

Here, $\theta > 0$ denotes a relaxation parameter which needs to be chosen appropriately. The iteration is done in the space of the nodal values of the pressure component. To balance possible irregularities of the computational grid $\mathcal{T}_h$ (anisotropic cells, heterogeneous grid spacing), it is advisable to precondition the system with the *mass matrix* of the trial space of the pressure,

$$M := M_Q = \left(\left(\chi_h^j, \chi_h^i\right)\right)_{i,j=1}^{N_Q}.$$

This leads to the so called *modified Uzawa algorithm*:

Algorithm 7.2 (Uzawa (modified)) Given an initial value y^0. For $l \in \mathbb{N}$ compute:

1.
$$Ax^l = b - By^{l-1}$$

2.
$$My^l = My^{l-1} + \theta B^\top x^l$$

Until convergence is reached.

Each iteration requires to invert the stiffness matrix A as the most costly operation. Furthermore, the mass matrix M needs to be inverted. Both can be accomplished in an efficient way with, e.g., the (preconditioned) CG or multigrid algorithms. By eliminating x^l (inserting the first step into the second one), we obtain the Uzawa algorithm in form of a fixed point iteration:

$$y^l = \left(I - \theta_l M^{-1}\Sigma\right) y^{l-1} + \theta_l M^{-1} B^\top A^{-1} b. \tag{7.2}$$

Therefore, the Uzawa algorithm is a damped Richardson iteration for solving the Schur complement equation. It holds the following result which is obtained by the application of Banach's fixed point theorem:

Theorem 7.3 (Uzawa algorithm) *Assume that the condition*

$$\theta_l \leq \frac{1-q}{\|M^{-1}\Sigma\|}, \quad q \in (0,1) \tag{7.3}$$

holds. Then Uzawa's algorithm converges to the solution (x, y) of the saddle-point problem (7.1). Furthermore, we have the following estimate on the error:

$$\left\|y^l - y\right\| \leq \frac{q^l}{1-q} \left\|y^0 - y\right\|.$$

Proof The exact solution y fulfils the fixed point equation

$$y = \left(I - \theta_l M^{-1}\Sigma\right) y + \theta_l M^{-1} B^\top A^{-1} b.$$

Substracting (7.2) yields the following estimate on the error $e^l := y - y^l$:

$$\left\|e^l\right\| \leq \left\|I - \theta_l M^{-1}\Sigma\right\| \cdot \left\|e^{l-1}\right\|.$$

This implies

$$\left\|e^l\right\| \leq \left|1 - \theta_l \left\|M^{-1}\Sigma\right\|\right| \cdot \left\|e^{l-1}\right\| \overset{(7.3)}{\leq} q \left\|e^{l-1}\right\|.$$

The Theorem now follows by the application of Banach's fixed point theorem. □

Remark 7.4

1. It can be shown, that $\left\|M^{-1}\Sigma\right\|$ is uniform for all grid spacings h. Therefore, the Uzawa algorithm converges with an order that is independent of h, if we additionally assume, that θ_l is small enough.

2. The rate of convergence can be optimized by choosing θ_l variable in each iteration. Then the Uzawa algorithm corresponds to the *method of steepest descent* for solving the Schur complement equation. We omit the detailed discussing and directly proceed to the more efficient CG method.

7.1.1.2 Conjugate gradient (CG) method

Similarly to the modified Uzawa algorithm, the system is preconditioned by the pressure mass matrix M_Q, i.e.: The CG method is applied as *preconditioned CG (PCG) method* to the modified system

$$M^{-1}\Sigma y = M^{-1}B^\top A^{-1}b.$$

This is equivalent to the application of the CG algorithm to symmetric and (up to mean value computation) positive definite matrix

$$M^{-\frac{1}{2}}\Sigma M^{-\frac{1}{2}}.$$

The spectral condition number

$$\kappa := \frac{\lambda_{\max}\left(M^{-1}\Sigma\right)}{\lambda_{\min}\left(M^{-1}\Sigma\right)}$$

determines the rate of convergence according to

$$\left\|y^l - y\right\| \leq \kappa \left(\frac{1-\kappa^{-\frac{1}{2}}}{1+\kappa^{-\frac{1}{2}}}\right)^l \left\|y^0 - y\right\|, \quad l \in \mathbb{N},$$

where y^0 is the initial guess of the iteration.

Theorem 7.5 (Schur complement) *The following estimate holds for the Schur complement* $\Sigma = B^\top A^{-1}B$*:*

$$\kappa\left(M^{-1}\Sigma\right) \leq c_0^{-1}\beta^{-2},$$

where $\beta > 0$ *denotes the constant from* (3.7) *of the used finite element discretization* X_h/Q_h*. In the conforming case* $X_h \subset X$ *it holds* $c_0 = 1$ *and* $c_0 = 4$ *in the general case.*

Proof Cf., e.g., [10]. □

The saddle-point problem (7.1) can also be solved by a multigrid algorithm. In this case, special attention has to be paid to the *smoother*. We will discuss that aspect in the context of the more general situation of the linearized Navier-Stokes equations.

7.2 Solution of the stationary Navier-Stokes equations

The solution of the *nonlinear* Navier-Stokes equations necessitates the application of an iterative method, e.g., Newton or Richardson iterations. In the case of dominating nonlinearities, i.e., dominance of the convective terms (and therefore high Reynolds numbers), some critical aspects need to be considered and respected.

7.2.1 Discretization of the convective terms

Consider the stationary Navier-Stokes equations in dimensionless form

$$\begin{cases} -\nu\Delta\mathbf{v} + (\mathbf{v}\cdot\nabla)\,\mathbf{v} + \nabla p & = f \\ \operatorname{div}\,\mathbf{v} & = 0 \end{cases}$$

PICTURE!!!

Figure 7.1: Solution of one-dimensional model problem

on some domain $\Omega \subset \mathbb{R}^d$, $(d = 2, 3)$ with sufficiently smooth boundary $\partial\Omega$. For the sake of simplicity, we restrict the following discussion to the case of *homogeneous* Dirichlet boundary conditions on the whole boundary $\partial\Omega$.

If we apply a conforming finite element discretization which fulfils (3.7), the variational formulation yields the following discrete contribution of the convective term:

$$((\mathbf{v} \cdot \nabla)\,\mathbf{v}, \psi) \approx ((\mathbf{v}_h \cdot \nabla)\,\mathbf{v}_h, \psi_h)\,.$$

This term mainly contributes to the off-diagonals of the system matrix, which consequently looses its definiteness. On coarser grids, this leads in turn to non-physical oscillations in the solution as well as the breakdown of the classical iterative solution methods, e.g., Jacobi, Gauß-Seidel or SOR algorithms.

7.2.1.1 Stabilization in the one-dimensional case

We illustrate these statements for the case of a simple one-dimensional model. On $\Omega = I := (0, 1) \subset \mathbb{R}$ we consider the boundary value problem

$$-\varepsilon u''(x) + q(x)u'(x) = 0,\ x \in I, \quad u(0) = 1,\ u(1) = 0$$

(*Sturm-Liouville problem*). We assume, that $q(x) \equiv 1$ on I. Then the unique solution reads

$$u^\varepsilon(x) = \frac{e^{\frac{1}{\varepsilon}} - e^{\frac{x}{\varepsilon}}}{e^{\frac{1}{\varepsilon}} - 1}.$$

In the case $\varepsilon \ll 1$ it holds for $x = 1 - \delta$ and $\delta > \varepsilon$

$$u^\varepsilon(1-\delta) = \frac{e^{\frac{1}{\varepsilon}}}{\frac{1}{\varepsilon} - 1}\left(1 - e^{-\frac{\delta}{\varepsilon}}\right) \approx 1, \quad \sup_{x \in I} \left|u^{\varepsilon\prime\prime}(x)\right| \approx \varepsilon^{-2},$$

which justifies the term *boundary layer*. For $\varepsilon = 0$ we obtain the solution $u^0 \equiv 1$, which obviously does not fulfil the boundary condition on the right boundary $x = 1$.

The approximation of this problem with the usual (centralized) difference scheme and an equidistant grid spacing $h = \frac{1}{N+1}$

$$\begin{aligned} -u''(x) &\approx \frac{1}{h^2}\left(-y_{n-1} + 2y_n - y_{n+1}\right) && \text{(2nd order accurate)}, \\ u'(x) &\approx \frac{1}{2h}\left(y_{n+1} - y_{n-1}\right) && \text{(2nd order accurate)}, \end{aligned}$$

yields

$$-\left(\varepsilon + \frac{h}{2}\right) y_{n-1} + 2\varepsilon y_n - \left(\varepsilon - \frac{h}{2}\right) y_{n+1} = 0, \quad 1 \le n \le N, \quad y_0 = 1,\ y_{N+1} = 0.$$

The resulting coefficient matrix is consequently only diagonally dominant under the restrictive condition

$$h \le 2\varepsilon.$$

For $h > 2\varepsilon$ the discrete solution incorporates non-physical oscillations. To verify this statement, we make the following trial for the solution:

$$y_n = \lambda^n.$$

The possible values for λ are the roots of the quadratic equation

$$\lambda^2 + \frac{2\varepsilon}{\frac{h}{2} - \varepsilon}\lambda - \frac{\frac{h}{2} + \varepsilon}{\frac{h}{2} - \varepsilon} = 0.$$

Therefore, the solution has the general form

$$y_n = c_+\lambda_+^n + c_-\lambda_-^n.$$

Considering the boundary conditions $y_0 = 1$ and $y_{N+1} = 0$ yields the solution

$$y_n = \frac{\lambda_+^{N+1}\lambda_-^n - \lambda_-^{N+1}\lambda_+^n}{\lambda_+^{N+1} - \lambda_-^{N+1}}, \quad n = 0, \ldots, N+1.$$

In our case, the roots are given by

$$\lambda_+ = 1, \quad \lambda_1 = \frac{\varepsilon + \frac{h}{2}}{\varepsilon - \frac{h}{2}}.$$

For $\varepsilon \ll \frac{h}{2}$ it holds $\lambda_- \approx -1$. In this case, we obtain an oscillating solution:

$$y_n = \frac{\lambda_-^n - \lambda_-^{N+1}}{1 - \lambda_-^{N+1}}, \quad n = 0, \ldots, N+1,$$

which does not reflect the correct solution. There are different strategies to omit this defect which we depict in the following:

1. **Upwind discretization:** Instead of using the central difference quotient for the approximation of the first derivative of u, we can chose one of the one-sided quotients

   $$\Delta_h^+ u(x) = \frac{u(x+h) - u(x)}{h}, \quad \Delta_h^- u(x) = \frac{u(x) - u(x-h)}{h}.$$

   If we chose the *backward* or *upwind* difference quotient Δ_h^-, we respect the physical phenomenon of information transport in positive x direction. The difference equations read

   $$(-\varepsilon + h)y_{n-1} + (2\varepsilon + h)y_n - \varepsilon y_{n+1} = 0$$

   and the resulting system matrix is diagonally dominant for arbitrary $h > 0$. The trial $y_n = \lambda^n$ for the solution yields the roots

   $$\lambda_+ = \frac{\varepsilon + h}{\varepsilon}, \quad \lambda_- = -1.$$

   The critical root is $\lambda_+ > 0$ in this case and the discrete solution

   $$y_n = \frac{\lambda_+^{N+1} - \lambda_+^n}{\lambda_+^{N+1} - 1}$$

   does not incorporate non-physical oscillations.

   The one-sided difference quotient is only first order accurate, i.e., $\mathcal{O}(h)$, and, therefore, the overall algorithm is only first order accurate. This limits the precision of the approximation in regions where the solution is smooth, even if the grid spacing h is chosen sufficiently small according to $h \approx \varepsilon$ in the boundary layer.

2. **Artificial diffusion:** We discretize the first drivative $u'(x)$ with the central difference quotient and set the *diffusion coefficient* to a larger value $\varepsilon_h := \varepsilon + \delta h$. *This is a change in the physics of the modelled problem!!!* In case of our model problem with $q \equiv 1$, the modified discretization is stable if we chose $\delta \geq \frac{1}{2}$. But the boundary layer is significantly smeared over to the interval $[1 - \varepsilon_h, 1]$ and the global accuracy of the approximation is only $\mathcal{O}(h)$ due to the perturbation of the differential operator.

We have seen at the hand of this simple example that the *damping strategies* upwind discretization and artificial diffusion lead on the one hand to stable discretizations but on the other hand the order of the approximation is reduced to $\mathcal{O}(h)$. Consequently, the question about *higher order stabilization schemes* for the discretization of *transport terms* is not answered in a satisfactory manner yet. Approaches in this direction use higher order one-sided difference quotient (in the case of *upwinding*) or artificial diffusion terms of the form $\delta h^2 u^{(iv)}$. But the strong M matrix property can only be achieved with discretizations of first order. *This restriction can be circumvented by a modification of the artificial diffusion in the context of Galerkin methods with finite element discretizations.*

3. **Streamline diffusion:** We consider the special case $q \equiv 1$. This method modifies the usual variational formulation

$$\varepsilon \left(u', \varphi'\right) + \left(u' + \alpha u, \varphi\right) = \left(f, \varphi\right), \quad \forall \varphi \in H := H_0^1(I),$$

with some $\alpha \geq 0$, to

$$\varepsilon \left(u', \varphi'\right) + \left(u' + \alpha u, \varphi + \delta\varphi'\right) = \left(f, \varphi + \delta\varphi'\right), \quad \forall \varphi \in H := H_0^1(I),$$

where δ denotes a parameter function which depends on the grid parameter h. The resulting bilinear form

$$a_\delta(u, v) := \varepsilon \left(u', v'\right) + \left(u' + \alpha u, \varphi + \delta\varphi'\right)$$

is coercive,

$$a_\delta(v, v) \geq \|v\|_\delta^2, \quad v \in H,$$

with respect to the modified energy norm

$$\|v\|_\delta := \left(\varepsilon \left\|v'\right\|^2 + \left\|\delta^{\frac{1}{2}} v'\right\|^2 + \alpha \|v\|^2\right)^{\frac{1}{2}}.$$

We want to emphasize that, in general, the parameter δ is a function of x (piecewise constant on the triangulation $0 = x_0 < \ldots < x_{N+1} = 1$) and, therefore, has to be considered within the norm $\left\|\delta^{\frac{1}{2}} v'\right\|$. Then, the corresponding finite-element-Galerkin method (with *linear trial functions*) in the trial space $H_h \subset H$ reads

$$\text{Find} \quad u_h \in H_h : \quad a_\delta(u_h, \varphi_h) = l_\delta(\varphi_h) \quad \forall \varphi_h \in H_h,$$

with the modified right hand side

$$l_\delta(v) := \left(f, v + \delta v'\right).$$

It can be shown that, in the case of a smooth solution u (without boundary layer) or if the grid spacing in the boundary layer is sufficiently small, the order of the approximation with respect to the energy norm is $\mathcal{O}\left(h^{\frac{3}{2}}\right)$. Consequently, the simplest finite element method with streamline diffusion has a higher order than the upwinding-stabilized finite difference method.

7.2.1.2 Stabilization of the Navier-Stokes equations

The methods for the stabilization of the transport terms that we described for the one-dimensional case can be generalized to the multi-dimensional Navier-Stokes equations.

1. **Upwind stabilization:** It is possible to transfer the idea of the upwind stabilization from one-dimensional finite differences to multi-dimensional finite element methods. The construction is technically quite involved because we need to determine the upwind direction or, more precisely, the "upwind area". We do not discuss this technique in detail here and refer the interested reader to [10].

2. **Artificial diffusion:** This approach works analogously to the one-dimensional case and incorporates the same limitations, especially the modification of the underlying physics.

3. **Streamline diffusion/Residual based stabilization:** The basic idea of *streamline diffusion* is to introduce artificial diffusion only in the direction of transport or in the direction of streamlines, respectively, such that the full second order of the discretization is preserved. There is a general construction principle behind this approach which we will discuss after a simple example.

   We describe a simple variant for the stationary Navier-Stokes equations: Find $\mathbf{v}_h \in X_h$ and $p_h \in Q_h$ such that

   $$\begin{aligned}((\mathbf{v}_h \cdot \nabla)\mathbf{v}_h, \varphi_h) + \nu\,(\nabla \mathbf{v}_h, \nabla \varphi_h) - (p_h, \operatorname{div}\,\varphi_h) + s_h\left(\{\mathbf{v}_h, p_h\}, \{\varphi_h, \chi_h\}\right) \\ = (f, \varphi_h) + r_h\left(\{\mathbf{v}_h, p_h\}, \{\varphi_h, \chi_h\}\right)\end{aligned}$$

   holds for all tuples $\{\varphi_h, \chi_h\} \in X_h \times Q_h$, where

   $$s_h\left(\{\mathbf{v}_h, p_h\}, \{\varphi_h, \chi_h\}\right) = \sum_{T \in \mathcal{T}_h} \delta_T \left\{(\nabla p_h + (\mathbf{v}_h \cdot \nabla)\mathbf{v}_h, \nabla \chi_h + \bar{\mathbf{v}}_h \nabla \varphi_h)_T + (\operatorname{div}\,\mathbf{v}_h, \operatorname{div}\,\varphi_h)_T\right\},$$
   $$r_h\left(\{\mathbf{v}_h, p_h\}, \{\varphi_h, \chi_h\}\right) = \sum_{T \in \mathcal{T}_h} \delta_T \left(f + \nu \Delta \mathbf{v}_h, \nabla \chi_h + \bar{\mathbf{v}}_h \nabla \varphi_h\right)_T$$

   with some reference velocity $\bar{\mathbf{v}}_h$. The stabilization parameters δ_T are chosen according to

   $$\delta_T = \min\left\{\frac{h_T^2}{\nu}, \frac{h_T}{\|\bar{\mathbf{v}}_h\|}\right\}.$$

   This choice of the stabilization has several special properties:

   a) $(\nabla p_h, \nabla \chi_h)_T$: Stabilizes the pressure for the conforming Q_1/Q_1 element such that (3.7) is fulfilled (with a modified numerator).

   b) $((\mathbf{v}_h \cdot \nabla)\mathbf{v}_h, \bar{\mathbf{v}}_h \nabla \varphi_h)_T$: Stabilizes the transport operator (streamline diffusion).

   c) $(\operatorname{div}\,\mathbf{v}_h, \operatorname{div}\,\varphi_h)_T$: Improves/stabilizes the conservation of mass (incompressibility).

   The other terms are only needed in order to preserve the second order accuracy of the discretization. A theoretical analysis shows that this kind of stabilization in the context of a Galerkin methods leads in fact to an improvement over the upwind stabilization: the discretization error on the Q_1/Q_1 element converges with $\mathcal{O}\left(h^{\frac{3}{2}}\right)$.

   General construction principle: The general principle behind the streamline diffusion method are the *residual based stabilization methods*. To motivate this approach, consider

a linear differential operator L and some right hand side function f. We rewrite the task of finding a solution u of the equation

$$Lu = f$$

in form of a minimization problem: Find u such that

$$u = \text{argmin}_v \left\| Lv - f \right\|^2 .$$

It can be shown that this is equivalent to compute u as the solution of the variational equation

$$(Lu - f, L\varphi) = 0 \quad \forall \varphi,$$

where u, φ are elements of some appropriate function space. This Galerkin method is known as *Galerkin-least squares (GLS) method*. Let us consider the last equation in the form

$$(Lu, L\varphi) = (f, L\varphi) \quad \forall \varphi.$$

The bilinear form

$$a_{GLS}(u, v) := (Lu, Lv)$$

is always *symmetric and coercive* which means that the matrix resulting from a finite element discretization is always *symmetric and positive definite*. This sounds like the "ideal" variational formulation but it has two severe disadvantages:

a) If L is a differential operator of the order $k \in \mathbb{N}$, we need to construct finite elements which are H^k conform. In the case $k = 1$ we can simply use our standard Lagrange finite element, but even for a second order differential operator like $L = -\Delta$ we need finite elements with higher global regularity. The construction of such elements is not trivial at all!

b) The matrix resulting from the bilinear form

$$a_{GLS}(u, v) := (Lu, Lv)$$

has a condition number that is squared compared to the "standard" variational formulation resulting from the approach

$$(Lu, \varphi) .$$

Due to these reasons, it is not advisable to use this kind of variational formulation for the discretization of a PDE except in very special or hard cases. But we can use this formulation *elementwise* for the stabilization of our discretization based on the "usual" variational formulation: Let $R(u) := Lu - f$ denote the residual of our PDE under consideration and $a(\cdot, \cdot)$ the bilinear form associated with L. Then the stabilized formulation reads: Find u_h such that

$$a(u_h, \varphi_h) - (f, \varphi) + \sum_{T \in \mathcal{T}_h} (R(u_h), \delta_T L\varphi_h)_T = 0 \quad \forall \varphi_h.$$

This is known as *Galerkin-least squares stabilization*. The advantage of the formulation is that it is strictly consistent with our PDE because the exact solution u fulfils this equations.

The evaluation of the the operator $L\varphi_h$ can be computationally expensive. So it is possible to generalize the GLS stabilization to the form

$$a(u_h, \varphi_h) - (f, \varphi) + \sum_{T \in \mathcal{T}_h} (R(u_h), \delta_T W \varphi_h)_T = 0 \quad \forall \varphi_h$$

with some *weighting operator* W. Two common choices in the context of convection-dominated problems are

- $W = L$: GLS stabilization
- $W = \bar{\mathbf{v}}_h \nabla \varphi_h$: The streamline-upwind Petrov-Galerkin (SUPG) stabilization operator.

In the case of a nonlinear operator $N(u)$, the weighting operator W is chosen as a linearization of N which incorporates the features that need to be stabilized.

So, the above simple variant of a residual based stabilization uses the GLS stabilization for the continuity equation and the operator

$$\nabla \chi_h + \bar{\mathbf{v}}_h \nabla \varphi_h$$

for the stabilization of the momentum equation which is the sum of pressure stabilization and SUPG stabilization.

Remark 7.6 The operator W can be interpreted as considering the contained parts of the differential operator L in a GLS manner because the part

$$(W(u_h), \delta_T W \varphi_h)_T$$

of the "full" stabilization operator

$$(R(u_h), \delta_T W \varphi_h)_T$$

is a GLS formulation for W.

4. **Local projection stabilization:** Streamline diffusion has significant disadvantages from a practical point of view. Its usage in conjunction with a Neumann outflow boundary condition ("do nothing" boundary condition) in channel flows leads to a non-physical flow behaviour ("numerical" boundary layer) at the outflow boundary: streamlines are flexed outwards. Furthermore, the evaluation of the additional terms can become expensive, especially in three-dimensional flow problems.

   These problems can be avoided by the application of the so called *local projection stabilization*. Here, the stabilization forms $r_h = 0$ and

   $$s_h := (\nabla (p_h - \pi_{2h} p_h), \delta_h \nabla (\chi_h - \pi_{2h} \chi_h)) + ((\bar{\mathbf{v}}_h \cdot \nabla)(\mathbf{v}_h - \pi_{2h} \mathbf{v}_h), \delta_h (\bar{\mathbf{v}} \cdot \nabla)(\varphi_h - \pi_{2h} \varphi_h))$$

   are used. π_{2h} is a projection or interpolation to the trial space L_{2h} which is defined on a coarser triangulation $\mathcal{T}_{2h}$. The resulting method is of second order, the evaluation of the corresponding system matrices relatively cheap and the error in consistency at an outflow boundary is avoided. Again, the parameter δ_h is chosen cellwise in dependence on the local convection and diffusion.

5. **Discontinuous Galerkin (DG) method:** The application of a discontinuous Galerkin discretization instead of a "standard" continuous Galerkin method has also a stabilizing effect but is not discussed here. We refer the interested reader to the literature.

7.2.2 Linearization

As we have already mentioned above, we need to *linearize* the *nonlinear* Navier-Stokes equations in order to obtain a discrete solution with the Newton or Richardson iterative methods, for example. We describe the various linearization techniques on the continuous level. Starting with some initial value $\mathbf{v}^0 \in X$ we determine a series of iterates $\left\{\mathbf{v}^l, p^l\right\} \in X \times Q$ for $l \in \mathbb{N}$ which converges to the exact solution $\{\mathbf{v}, p\}$.

7.2.2.1 Stokes linearization

In the case of small Reynolds numbers

$$Re = \frac{1}{\nu} \ll 1$$

(viscous liquids, small velocities, small geometrical dimensions), the simple *Stokes linearization*

$$\begin{cases} -\nu\Delta\mathbf{v}^l + \nabla p^l & = f - \left(\mathbf{v}^{l-1} \cdot \nabla\right)\mathbf{v}^{l-1} \\ \text{div } \mathbf{v}^l & = 0 \end{cases}$$

is often sufficient. The nonlinearity $(\mathbf{v} \cdot \nabla)\,\mathbf{v}$ is treated fully explicit and in each iteration the symmetric and positive definite Stokes operator has to be inverted.

7.2.2.2 Oseen linearization

If the Reynolds numbers are larger, the so called *Oseen linearization* is used:

$$\begin{cases} -\nu\Delta\mathbf{v}^l + (\bar{\mathbf{v}} \cdot \nabla)\,\mathbf{v}^l + \nabla p^l & = f \\ \text{div } \mathbf{v}^l & = 0 \end{cases}.$$

Here, $\bar{\mathbf{v}} \approx \mathbf{v}^l$ is an appropriate approximation, e.g., in the simplest case $\bar{\mathbf{v}} := \mathbf{v}^{l-1}$ (*constant* extrapolation), but also higher order approximations are possible (e.g., with *linear* extrapolation). The operator which has to be inverted in each iteration is a *non-symmetric* convection-diffusion operator. In order to improve the convergence, this *functional iteration* should be done in form of a defect correction iteration: A correction is computed with the defect

$$d^{l-1} := f - \nu\Delta\mathbf{v}^{l-1} - \left(\mathbf{v}^{l-1} \cdot \nabla\right)\mathbf{v}^{l-1} - \nabla p^{l-1}$$

from the equations

$$\begin{cases} -\nu\Delta\mathbf{w}^l + (\bar{\mathbf{v}} \cdot \nabla)\,\mathbf{w}^l + \nabla q^l & = d^{l-1} \\ \text{div } \mathbf{w}^l & = 0 \end{cases} \tag{7.4}$$

and afterwards the (approximate) solution is updated to

$$\begin{aligned} \mathbf{v}^l &:= \mathbf{v}^{l-1} + \lambda_l \mathbf{w}^l, \\ p^l &:= p^{l-1} + \lambda_l q^l. \end{aligned}$$

The parameter λ_l is used to damp the iteration and needs to be computed adaptively. To save computational costs, the nonsymmetric term $(\bar{\mathbf{v}} \cdot \nabla)\,\mathbf{w}^l$ is often not updated in every iteration; in extreme cases, a fixed approximation $\bar{\mathbf{v}} \approx \mathbf{v}^0$ is used.

7.2.2.3 Newton linearization

The iteration (7.4) can we rewritten in the form

$$\begin{cases} -\nu\Delta\mathbf{v}^l + \left(\mathbf{v}^{l-1}\cdot\nabla\right)\mathbf{v}^l + \left(\mathbf{v}^l\cdot\nabla\right)\mathbf{v}^{l-1} + \nabla p^l & = f - \left(\mathbf{v}^{l-1}\cdot\nabla\right)\mathbf{v}^{l-1} \\ \operatorname{div}\mathbf{v}^l & = 0 \end{cases}.$$

This is equivalent to the classical *Newton method* to compute a zero point of the system of equations

$$\begin{cases} -\nu\Delta\mathbf{v} + \left(\mathbf{v}\cdot\nabla\right)\mathbf{v} + \nabla p - f & = 0 \\ \operatorname{div}\mathbf{v} & = 0 \end{cases}.$$

7.2.3 Algebraic solution of the linearized problems

We consider finite element discretizations, possibly with pressure stabilization. With the natural nodal basis functions of the trial spaces X_h and Q_h the discrete , e.g., Oseen equations (7.4) can be written in block matrix form

$$\mathcal{A}(\bar{\mathbf{v}})\xi = \begin{pmatrix} A(\bar{\mathbf{v}}) & B \\ -B^\top & C \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ c \end{pmatrix} \tag{7.5}$$

for the velocity's and pressure's nodal values vectors $\xi = (x, y)^\top$ with the corresponding matrices and right hand side vectors

$$\begin{aligned} A(\bar{\mathbf{v}}) &= \left(\left(\nu\nabla\psi_h^j, \nabla\psi_h^i\right) + \left(\left(\bar{\mathbf{v}}\cdot\nabla\right)\psi_h^j, \psi_h^i\right) + \ldots\right)_{i,j=1}^{N_X}, \\ B &= \left(-\left(\chi_h^j, \operatorname{div}\psi_h^i\right)\right)_{i,j=1}^{N_X,N_Q}, \\ C &= \left(\left(\nabla\chi_h^j, \delta\nabla\chi_h^i\right) + \ldots\right)_{i,j=1}^{N_Q}, \\ b &= \left(\left(f, \psi_h^i\right)\right)_{i=1}^{N_X}, \\ c &= \left(\left(g, \chi_h^i\right)\right)_{j=1}^{N_Q}. \end{aligned}$$

C denotes the matrix which originates from a possible pressure stabilization; otherwise, $C = 0$. The stabilization of the transport terms is incorporated in the matrix $A(\bar{\mathbf{v}})$, but is not explicitly stated in order to simplify the notation. The matrix $A(\bar{\mathbf{v}})$ is nonsymmetric but (under natural conditions) regular. Therefore, we can compute a Schur complement again and use the Uzawa algorithms: We eliminate the x component from the system

$$\begin{aligned} x &= A(\bar{\mathbf{v}})^{-1}(By + b), \\ \left(C - B^\top A(\bar{\mathbf{v}})^{-1}B\right)y &= c - B^\top A(\bar{\mathbf{v}})^{-1}b. \end{aligned}$$

Again, we have a block triangular decomposition with the *Schur complement* $\Sigma := C - B^\top A(\bar{\mathbf{v}})^{-1}B$,

$$\mathcal{A}(\bar{\mathbf{v}}) = \begin{pmatrix} A(\bar{\mathbf{v}}) & B \\ -B^\top & C \end{pmatrix} = \begin{pmatrix} A(\bar{\mathbf{v}}) & 0 \\ -B^\top & \Sigma \end{pmatrix}\begin{pmatrix} I & A(\bar{\mathbf{v}})^{-1}B \\ 0 & I \end{pmatrix}.$$

Consequently, the modified Uzawa algorithm reads

$$\begin{aligned} A(\bar{\mathbf{v}})x^l &= b - By^{l-1}, \\ My^l &= My^{l-1} + \theta_l\left(B^\top x^l + c\right). \end{aligned}$$

In each iteration the computationally most intensive operations are the inversions of the matrices $A(\bar{\mathbf{v}})$ and M which can be done with the GMRES or a multigrid methods, for example. It can again be shown, that the Uzawa method converges independently of the underlying computational grid.

Due to the instabilities caused by the transport term, the convergence slows down more and more with increasing Reynolds numbers such that the Uzawa method is inefficient quite soon. An alternative is provided by the multigrid method which we describe next.

7.2.3.1 Multigrid method

The main ideas behind the *multigrid algorithm* are the fast reduction of *high-frequency* error contributions (*smoothing*) via *cheap* relaxation method (e.g. Jacobi or Gauß-Seidel iterations) on a fine grid and the reduction of the remaining *smooth*, low-frequency error contribution via a defect correction on coarser grids (*coarse grid correction*). We describe this approach shortly:

The multigrid iteration uses a hierarchy of finite element subspaces

$$V_0 \subset V_1 \subset \ldots \subset V_L,$$

which can be obtained by a systematic refinement process. The connections between these spaces are established via *prolongation* operations

$$P_{l-1}^{l} : V_{l-1} \to V_l$$

and *restriction* operations

$$R_{l}^{l-1} : V_l \to V_{l-1}.$$

In the context of finite elements, the natural choices are

$$P_{l-1}^{l} : \text{ natural embedding} \quad \text{and} \quad R_{l}^{l-1} : \ L^2 \text{ projection.}$$

The main part of a multigrid method are the smoother iterations

$$S_l : V_l \to V_l$$

on the different grid levels $0 \leq l \leq L$ ($l = 0$ corresponds to the coarsest grid and $l = L$ to the finest one). The multigrid iteration

$$\mathcal{M}\xi = \mathcal{M}\left(l, z^0, \xi\right)$$

on grid level l with initial value z^0, m_1 *pre-smoothing* and m_2 *post-smoothing* steps is recursively defined as follows:

Algorithm 7.7 (Multigrid method) For $l = 0$ the multigrid algorithm is given by an exact solver, i.e.,

$$\mathcal{M}\left(0, z^0, \xi\right) := \mathcal{A}_0(\bar{\mathbf{v}})^{-1}\xi.$$

For $l \geq 1$, the following iteration is performed:

1. m_1 times pre-smoothing:
$$z_1 := S_l^{m_1} z^0.$$

2. Residual on level l:
$$r^l := \xi - \mathcal{A}_l(\bar{\mathbf{v}})z^0.$$

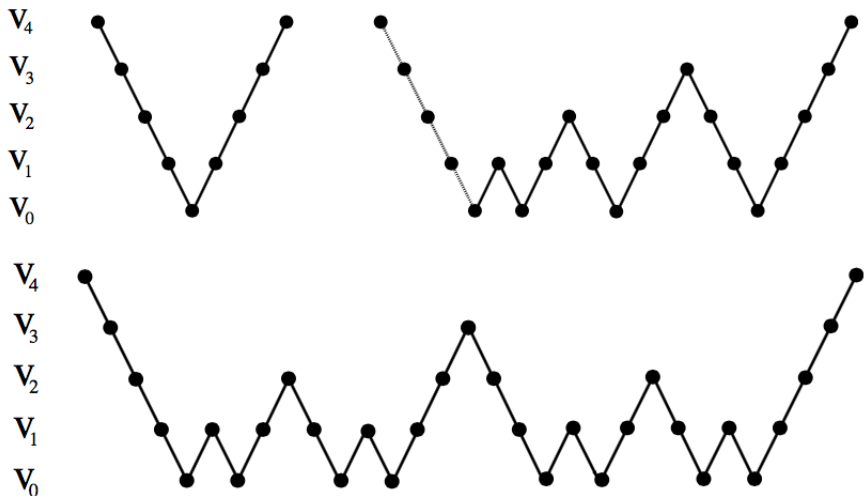

Figure 7.2: Schematic representation of the V cycle (top left), F cycle (top right) and W cycle (bottom)

3. Restriction to level $l-1$:
$$\tilde{r}^{l-1} := R_l^{l-1} r^l.$$

4. Coarse grid correction with initial value $q^0 := 0$:
$$q := \mathcal{M}\left(l-1, q^0, \tilde{r}^{l-1}\right).$$

5. Prolongation to level l:
$$z^2 := z^1 + P_{l-1}^l q.$$

6. m_2 times post-smoothing:
$$\mathcal{M}\left(l, z^0, \xi\right) := S_l^{m_2} z^2.$$

If the multigrid algorithm is applied γ times on each grid level, then the resulting scheme is called

- *V cycle* in the case $\gamma = 1$ and
- *W cycle* in the case $\gamma = 2$.

The variants with $\gamma \geq 3$ are computationally too expensive and are not used in practice. If the multigrid algorithm is only used as a preconditioner for a robust iteration, e.g., the GMRES or CG methods, a V cycle is sufficient in most cases. If the multigrid algorithm is the primary solver, the W cycle is more robust, especially for non-symmetric problems, and is preferred. The so called *F cycle* is an appealing compromise between V and W cycle.

The construction of a multigrid method for the solution of the saddle-point problem (7.5) is a delicate task. Especially, the choice of the smoothing iteration has to be taken with care because the usual fixed point iterations like Jacobi or Gauß-Seidel methods fail in this case. This problem can be tackled in several ways:

1. **Damped Jacobi smoother:** In the case $C \neq 0$ the matrix $\mathcal{A}(\bar{\mathbf{v}})$ is only weakly-definite which enables the application of standard methods like damped Jacobi iterations. But the resulting algorithms are not very robust and the choice of the iteration parameters has to be taken with care. Furthermore, for large Reynolds numbers this method becomes inefficient soon and does not converge anymore.

2. **Block Gauß-Seidel smoother ("Vanka" smoother):** A simple, but very successful, smoother is obtained for the matrix $\mathcal{A}(\bar{\mathbf{v}})$ by cellwise blocking of the degrees of freedom of all physical variables within one global Gauß-Seidel iteration. This was originally proposed for a finite difference approximation. We describe the realisation of this idea shortly: the degrees of freedom of velocity and pressure variables on one cell T or on a group of cells are combined, i.e., numbered consecutively, and the corresponding element-wise system matrices are marked by the index "loc". These local degrees of freedom are now iterated simultaneously within one Gauß-Seidel iteration:

$$S_{loc}(\bar{\mathbf{v}})\mathbf{v}_{loc}^k + B_{loc}p_{loc}^k = \text{"known"}, \quad B_{loc}^\top \mathbf{v}_{loc}^k = \text{"known"},$$

where $S_{loc} := A_{loc}(\bar{\mathbf{v}})$. This iteration runs over all blocks of cells. To increase the robustness of this method, the iteration is typically damped.

Remark 7.8 (Multigrid iteration as preconditioner) Let $\mathcal{A}$ be the system matrix of a finite element discretization of the linearized Navier-Stokes equations. The theory of multigrid methods is very well developed for scalar problems, whereas the situation for systems of equations is not so clear. We know that the application of multigrid iterations as a preconditioner within an outer iteration (e.g., GMRES or CG) underlies weaker restrictions than the usage as a stand-alone solver. Let $\mathcal{M}$ denote the operator of one multigrid step, then it is sufficient to have an upper bound on the condition number of the product $\mathcal{M}\mathcal{A}(\bar{\mathbf{v}})$, whereas in the stand-alone case the eigenvalues of the iteration matrix $\mathcal{B} := \mathcal{I} - \mathcal{M}\mathcal{A}(\bar{\mathbf{v}})$ need to be bounded uniformly away from one. For these reasons, in the context of flow problems the multigrid algorithm is only used as a preconditioner within a robust outer GMRES iteration. The multigrid operator $\mathcal{M}$ is interpreted as an approximate inverse $\mathcal{M} \approx \mathcal{A}(\bar{\mathbf{v}})^{-1}$. Of course it is not necessary, to compute this matrix explicitly: it is sufficient, to evaluate matrix-vector products $\mathcal{M}\xi$, i.e., to execute the multigrid process for a given and fixed right hand side.

7.3 Solution of the instationary Navier-Stokes equations

7.3.1 Time-stepping schemes

The system of the instationary Navier-Stokes equations

$$\begin{cases} \partial_t \mathbf{v} - \nu\Delta\mathbf{v} + (\mathbf{v}\cdot\nabla)\,\mathbf{v} + \nabla p & = f \\ \operatorname{div}\,\mathbf{v} & = 0 \end{cases}$$

is, in the terminology of ordinary differential equations, a so called *differential-algebraic system of equation (DAE)*. The *algebraic* constraint

$$\operatorname{div}\,\mathbf{v} = 0$$

determines a linear manifold in the so called *phase space* $H = H_0^1(\Omega)^d$, where the dynamics are taking place.

To solve this problem numerically, we proceed according to *Rothe's method*, i.e., we first discretize in time and afterwards in every time-step the occuring quasi-stationary problem is discretized in space.

Remark 7.9 The discretization strategy, where we first discretize in space and afterwards in time, is called *methods of lines*.

Let

$$0 = t_0 < t_1 < \ldots < t_m < \ldots < t_M = T$$

discrete points in time,

$$k_m := t_m - t_{m-1}$$

the time-step sizes and

$$k := \sup_m k_m.$$

To determine the approximations $\{\mathbf{v}^m, p^m\} \approx \{\mathbf{v}(\cdot, t_m), p(\cdot, t_m)\}$, we first consider the so called *one-step θ scheme* with the parameter $\theta \in [0, 1]$:

$$\begin{aligned}\theta f^m + (1-\theta) f^{m-1} = \frac{1}{k_m}\left\{\mathbf{v}^m - \mathbf{v}^{m-1}\right\} - \nu\Delta\left\{\theta\mathbf{v}^m + (1-\theta)\mathbf{v}^{m-1}\right\} \\ + \theta\left(\mathbf{v}^m \cdot \nabla\right)\mathbf{v}^m + (1-\theta)\left(\mathbf{v}^{m-1} \cdot \nabla\right)\mathbf{v}^{m-1} \\ + \nabla\left\{\theta p^m + (1-\theta)p^{m-1}\right\},\end{aligned}$$

where $f^m := f(t_m)$ and the constraint

$$\operatorname{div} \mathbf{v}^m = 0$$

has to be fulfilled. Depending on the choice of θ, we obtain the following time-stepping schemes:

- $\theta = 1$: implicit Euler

  $$\begin{pmatrix} \frac{1}{k_m} I + A + \hat{A} & B \\ -B^\top & 0 \end{pmatrix} \begin{pmatrix} v_m \\ p_m \end{pmatrix} = \begin{pmatrix} R_{m-1} \\ 0 \end{pmatrix},$$

  where I is the identity/mass matrix, A the stiffness matrix, $\hat{A}$ the matrix resulting from the nonlinearity and B as defined before. If $k_m \to 0$, then $\frac{1}{k_m} I + A + \hat{A} \approx \frac{1}{k_m} I$. The resulting matrix

  $$\begin{pmatrix} \frac{1}{k_m} I & B \\ -B^\top & 0 \end{pmatrix}$$

  is regular and well-conditioned. In the case $k_m \to \infty$, then $\frac{1}{k_m} I + A + \hat{A} \approx A + \hat{A}$. The resulting matrix

  $$\begin{pmatrix} A + \hat{A} & B \\ -B^\top & 0 \end{pmatrix}$$

  is regular but the condition number of A is on the order $\mathcal{O}\left(h^{-2}\right)$ and the matrix therefore becomes more and more ill-conditioned if $h \to 0$.

- $\theta = 0$: explicit Euler

- $\theta = \frac{1}{2}$: Crank-Nicolson

One time-step $y^{m-1} \to y^m = R(kq)y^{m-1}$ assumin the test equation

$$y'(t) = qy(t) \quad t \geq 0.$$

The solution of this test equation is clearly

$$y(t) = e^{qt}y(0).$$

The *amplification factor* $R(kq)$ fulfils the following approximation:

$$R(kq) = e^{kq} + \mathcal{O}\left(k^{r+1}\right).$$

- A stability:
$$|R(\lambda)| \leq 1.$$
- Global stability:
$$\lim_{Re(\lambda)\to\infty} |R(\lambda)| \leq 1 - \mathcal{O}(k).$$
- Strong A stability:
$$\lim_{Re(\lambda)\to\infty} |R(\lambda)| \leq 1 - \delta < 1.$$
- Weak dissipation:
$$|R(\lambda)| = 1 - \mathcal{O}\left(|Im(\lambda)|\right), \quad Re(\lambda) \to 0.$$

7.3.2 Projection methods

7.3.2.1 Chorin method

This method is based on the implicit Euler scheme:

$$\begin{cases} \frac{1}{k_m}\left\{\mathbf{v}^m - \mathbf{v}^{m-1}\right\} - \nu\Delta\mathbf{v}^m + \left(\mathbf{v}^m \cdot \nabla\right)\mathbf{v}^m + \nabla p^m & = f^m \\ \text{div } \mathbf{v}^m & = 0 \end{cases}$$

Based on the information of $\{\mathbf{v}^{m-1}, p^{m-1}\}$ we consider the following equation:

$$\frac{1}{k}\left\{\tilde{\mathbf{v}}^m - \mathbf{v}^{m-1}\right\} - \nu\Delta\tilde{\mathbf{v}}^m + \left(\tilde{\mathbf{v}}^m \cdot \nabla\right)\tilde{\mathbf{v}}^m = f^m.$$

Bibliography

[1] http://upload.wikimedia.org/wikipedia/commons/b/b1/Vehicle_crash_test_at_the_General_Motors_Vehicle_Safety_and_crash_Worthiness_Laboratory.jpg

[2] http://upload.wikimedia.org/wikipedia/commons/9/97/The_Earth_seen_from_Apollo_17.jpg

[3] http://www.nasa.gov/images/content/148709main_d4_testing_08.jpg

[4] http://upload.wikimedia.org/wikipedia/commons/f/fe/Airplane_vortex_edit.jpg

[5] http://upload.wikimedia.org/wikipedia/commons/2/2d/Aeroakustik-Windkanal-Messhalle.JPG

[6] http://upload.wikimedia.org/wikipedia/commons/6/67/Windpark_Bassens.jpg

[7] Gert Böhme: *Strömungsmechanik nichtnewtonscher Fluide*, B. G. Teubner Stuttgart/Leipzig/Wiesbaden, 2. Auflage, 2000

[8] Franco Brezzi, Michel Fortin: *Mixed and Hybrid Finite Element Methods*, Springer New-York, 1991

[9] William J. Layton: *Introduction to the Numerical Analysis of Incompressible Viscous Flows*, SIAM, Computational Science & Engineering, 2008

[10] Rolf Rannacher: *Numerische Mathematik 3 - Numerik von Problemen der Kontinuumsmechanik*, Lecture notes, Heidelberg University, 2008

Index

a posteriori error, 55
a priori error, 55
algorithm
    modified Uzawa, 62
    steepest descent, 63
    Uzawa, 62

boundary layer, 65
bubble function, 52

Clay prize, 44
coefficient
    diffusion, 67
complement
    Schur, 62
convergence
    super-, 59
criterion
    Fortin, 52
curl, 10

DAE, 75
damping strategy, 67
derivative
    directional, 10
    material, 16
    partial, 10
    total, 10
DG method, 70
differential-algebraic equations, 75
diffusion coefficient, 67
directional derivative, 10
Discontinuous Galerkin (DG), 70
discretization
    upwind, 66
divergence, 10
domain, 10
dual problem, 56

element
    MINI, 52
error
    a posteriori, 55
    a priori, 55

fluid
    Newtonian, 28
formulation
    mixed, 50
Fortin criterion, 52
function
    bubble, 52
    Green's, 56
    Lagrange, 47, 48
    weight, 56
functional
    Lagrange, 47, 48

Galerkin-least squares(GLS), 69
GLS, 69
gradient, 10
Green's function, 56

hydrostatic pressure, 27

inf-sup condition
    discrete, 33

Lagrange function, 47, 48
Lagrange functional, 47, 48
Laplace operator, 10
layer
    boundary, 65
least squares, 69
linearization
    Oseen, 71
    Stokes, 71
local projection stabilization (LPS), 70

material derivative, 16
method
    Discontinuous Galerkin (DG), 70
    line, 76
    Newton, 72
    Rothe's, 76
    steepest descent, 63

method of lines, 76
MINI element, 52
mixed formulation, 50
modified Uzawa algorithm, 62

nabla operator, 10
Navier-Stokes equations
    unsteady incompressible, 29
Newton method, 72
Newtonian fluid, 28

operator
    Laplace, 10
    nabla, 10
Oseen linearization, 71

partial derivative, 10
phase space, 75
pressure
    hydrostatic, 27
primal problem, 56
principal strain, 21
prize
    Clay, 44
problem
    dual, 56
    primal, 56
    saddle-point, 33
    Stokes, 33
    Sturm-Liouville, 65

refinement strategy, 56
residual based stabilization, 68
Reynolds number, 29
Rothe's method, 76

saddle-point problem, 33, 47, 48
Schur complement, 62
shear strain, 22
shear tensor, 27
shear viscosity, 29
solution
    strong, 43
    turbulent, 44
    weak, 43, 45
space
    phase, 75
spatial variable, 10
stabilization
    local projection, 70
    residual based, 68
steepest descent, 63
Stokes linearization, 71
Stokes problem, 33
strain, 21
    principal, 21
    shear, 22
    tensor, 21
    velocity, 23
strain tensor, 21
strategy
    damping, 67
    refinement, 56
Streamline-upwind Petrov-Galerkin, 70
strong solution, 43
Sturm-Liouville problem, 65
superconvergence, 59
SUPG, 70

tension, 50
tensor
    rate of deformation, 24
    shear, 27
    strain, 21
time variable, 10
total derivative, 10
turbulent solution, 44

upwind discretization, 66
Uzawa algorithm, 62
    modified, 62

variable
    spatial, 10
    time, 10
velocity
    strain, 23
viscosity
    shear, 29
    volume, 29
volume viscosity, 29

weak solution, 43, 45
weight function, 56