Hybrid Discontinuous Galerkin methods for solving incompressible flow problems

Diplomarbeit

zur Erlangung des akademischen Grades

Diplomingenieur

in der Studienrichtung

Computational Engineering Science

an der

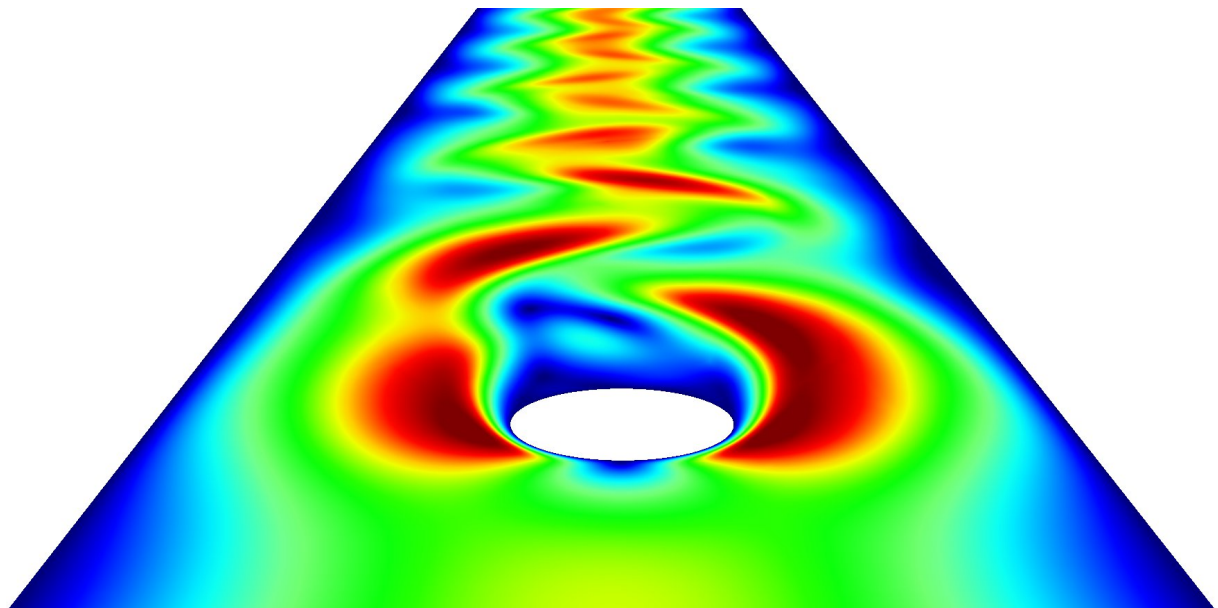Rheinisch-Westfälischen Technischen Hochschule Aachen

Autor:
Christoph Lehrenfeld

Betreuer:
Prof. Dr. Joachim Schöberl
Prof. Dr. Arnold Reusken

June 6, 2010

Abstract

This thesis deals with a higher order discretization of incompressible flow problems using the Hybrid Discontinuous Galerkin Method. It aims to introduce an appropriate and computationally efficient Hybrid Discontinous Galerkin formulation for the most important model problems of incompressible fluid flow, namely the convection diffusion equation and the incompressible Navier-Stokes equations.

The main contribution is the derivation, discussion and analysis of the Hybrid (exactly divergence-free) Discontinuous Galerkin method which results from the combination of the following three modern and promissing concepts in the field of (Discontinuous) Finite Element Methods:

The first of those concepts goes back to [CGL08] where the hybridization of Discontinuous Galerkin Methods is proposed resulting in the Hybrid Discontinuous Galerkin Method which by construction compensates for the computational disadvantages of Discontinuous Galerkin Method and thereby allows for a more efficient use its advantages. Linear systems of equations stemming from a Hybrid Discontinuous Galerkin formulation can be reduced to interface unknowns similar to the static condensation approach. This reduces the computational effort of Discontinuous Galerkin Method.

The second concept is motivated by [CKS05] where it was shown that Discontinuous Galerkin Methods for the Navier-Stokes equations should provide exactly divergence-free solutions in order to be stable and locally conservative. This leads to the idea of using divergence-conforming Finite Elements, which are normal-continuous across element interfaces, instead of using completely discontinous ones.

In [Zagl06] and [SZ05] a strategy for the construction of conforming higher order Finite Elements for H^1, $H(\mathrm{curl})$, $H(\mathrm{div})$ and L^2 based on the exact sequence property of those spaces is proposed and discussed. This approach leads - beside other advantages - to a natural separation of the $H(\mathrm{div})$-conforming Finite Elements into basis functions which are relevant for the discrete solution of the incompressible Navier-Stokes equations and those which can be neglected (a priori). This leads to a reduction of the number of unknowns.

Acknowledgements

Eidesstattliche Erklärung

Ich versichere hiermit, die vorgelegte Arbeit in dem gemeldeten Zeitraum ohne fremde Hilfe verfaßt und mich keiner anderen als der angegebenen Hilfsmittel und Quellen bedient zu haben.
Diese Diplomarbeit wurde bisher weder im In- noch im Ausland als Prüfungsarbeit vorgelegt.

Aachen, den June 6, 2010

(Vorname, Nachname)

Contents

Introduction

Motivation

The numerical solution of partial differential equations (PDE) related to (incompressible) flow problems is an important issue in many engineering applicitions. Facing the task of solving partial differential equations numerically, one has to choose between quite many different available methods. In Computational Fluid Dynamics the Finite Volume Method as well as the Standard (Continuous Galerkin) Finite Element Methods (CG FEM) are often used as they - in contrast to the Finite Difference Method - also work on unstructured meshes.

The concept of Standard (Continuous Galerkin) Finite Element Methods (CG FEM) is based upon the variational formulation of the PDE. An approximation to the PDE is achieved by replacing the infinite dimensional space in which the variational formulation is posed by a finite dimensional subspace which normally uses (element-) piecewise polynomials which are continuous across element interfaces. This allows for increasing the order of approximation quite easily also on unstructured meshes as only the polynomial degree has to be increased (given that the solution is sufficiently smooth). However, the CG FEM suffers from stability issues for convection-dominated flows and the lack of a conservation property. Finite Volume Methods (FVM) on the other hand, which are frequently used in engineering practice and are build upon the balance of in- and outfluxes on a control volumes, have neither of these problems. But here the problem is that providing higher order approximations with Finite Volume Methods is hardly possible.

A modern approach which combines the features and overcomes the draw-backs of both methods is the use of *Discontinuous Galerkin Finite Element Methods* (DG FEM). It is also based upon a variational formulation but this time the finite dimensional space is not taken to be a subspace of the infinite one. Instead, the finite one uses discontinuous functions, which are not allowed in the continuous setting. As the DG FEM approximation with piecewise constant functions coincides with the use of a Finite Volume Method the Finite Volume method is, in a sense, a subclass and thereby its conservation property can be inherited to the DG FEM. Although all stated problems of FVM and CG FEM can be solved with DG FEM, DG FEM has one (big) disadvantage:

It introduces considerably more unknowns as CG FEM or FVM. Especially linear systems arising from discretizations with DG FEM are considerably larger and less sparse as a lot of couplings between unknowns of different elements are introduced.

One approach to tackle this problem is the concept of hybridization which has already been used for discretization such as mixed methods (see [BF91]). The general concept which also allows the hybridization of DG FEM was proposed in [CGL08]. This approach leads to the *Hybrid Discontinuous Galerkin Finite Element Method* (HDG FEM) which is the major subject of this diploma thesis. As we will see, HDG FEM is actually a subset of DG FEM, but allows for a more efficient solution of linear systems stemming from the discretization.

For the incompressible Navier-Stokes equations it was shown in [CKS05] that DG FEM discretizations have to provide exactly divergence-free solutions to ensure a locally conservative, energy-stable and optimally convergent method. This motivates the use of divergence-conforming Discontinuous Galerkin Finite Elements and, to combine this with the aspects mentioned before, to transform it to a HDG FEM formulation. Furthermore the construction of higher order divergence-conforming Finite Elements possesses another potential. The construction of those Finite Elements can be based on the exact sequence property of the Sobolev spaces H^1, $H(\text{curl})$, $H(\text{div})$ and L^2 which allows for further improvement of the HDG FEM formulation for the incompressible Navier-Stokes equations. For this type of construction, a natural separation of the Finite Element space takes place. One subset of this Finite Element space can then be removed without consequences on the approximation of (strongly) divergence-free velocity fields. This reduces the total number of unknowns compared to other discretizations.

The above adressed issues are only problems in the spatial discretization and adding time integration with the method of lines seems to be clear. Here you normally use either fully implicit or explicit methods, like Runge-Kutta schemes. But the structure of convection diffusion equations - and the Navier-Stokes equations belong to this class of equations - can be considered for by semi-implicit time integration methods. Those combine the features of implicit and explicit methods resulting in a method which has to solve linear systems of equations which are far more easy to handle than those arising in fully implicit schemes but whose time step restriction - basically - only depends on the convection term of the PDE which is a mild condition compared to time step restrictions for fully explicit methods. As this approach is not as famous as the fully explicit and fully implicit counterparts, we will at least present higher order time integration schemes of this kind, which have been proposed in [ARW95] and [ARS97], the so called IMEX (implicit-explicit) schemes.

As two model problems on which we'll apply the above indicated strategies we consider the scalar *convection-diffusion equation* and the *incompressible Navier-Stokes equations*.

Organization of the work

The thesis is organized as follows:

Chapter 1 derives an HDG FEM formulation for the steady scalar convection diffusion equation by considering an HDG FEM formulation for each limit case, i.e. the pure convective one (linear transport equation) and the pure diffusive one (Poisson's equation), first and joining them afterwards. Comparisons and relations to other (e.g. Discontinuous or Mixed) Finite Element Methods are discussed and a rather complete a priori error analysis is carried out. In addition, numerical examples and the discussion of properties and computational aspects substantiate the potential of Hybrid Discontinuous Galerkin Methods.

Chapter 2 translates the discretization discussed in chapter 1 to the steady incompressible Navier-Stokes equations. Before the new divergence-conforming HDG formulation is introduced an extra section about the constructing of higher order $H(\text{div})$-conforming Finite Elements is given. This section includes the discussion of the construction proposed and discussed in [SZ05, Zagl06] which allows for the reduction of the number of unknowns. Starting from the Stokes problem the new discretization

of the steady incompressible Navier-Stokes equations is analyzed and discussed.

Chapter 3, which is about time integration (especially of semi-implicit type) for Navier-Stokes and convection diffusion equations, brings the diploma thesis to the conclusion. Therein the IMEX time integration method is derived and two different approaches to achieve higher order time integration of semi-implicit type are suggested.

The Appendix includes a summary of the notation used through out the thesis, some elementary inequality and the definition of some orthogonal polynomials. Furthermore two (more involved) proofs which were moved from chapter 2 can be found at the end of the appendix.

Requirements on the reader

The reader is presumed to have basic knowledge of the Finite Element Method including basic knowledge of functional analytic concepts. No knowledge of nonconforming Finite Elements like DG FEM is required.

Implementations

All implementations used for the numerical examples in chapter 1 were carried out with the open source software package `NETGEN/NGSolve`. `NETGEN` is an automatic mesh generator which is available at

$$\texttt{http://sourceforge.net/projects/netgen-mesher}$$

The Finite Element Library `NGSolve` which comes as an Add-On to `NETGEN` is also open source and available at

$$\texttt{http://sourceforge.net/projects/ngsolve}$$

It already includes the proposed HDG discretization for the scalar problem.
Implementations of the Stokes and Navier-Stokes discretization discussed in chapter 2 and also the time integration schemes of chapter 3 can be found in the `NGSolve-Add-On` package `ngsflow` which is developed by Joachim Schöberl and the author. It is available under

$$\texttt{http://sourceforge.net/projects/ngsflow}$$

and is able to solve the steady and unsteady incompressible Navier-Stokes problem on unstructured (hybrid) 2D and 3D meshes with arbitrary polynomial degree.

Chapter 1

The Hybrid Discontinuous Galerkin method for steady scalar convection diffusion problems

1.1 The scalar convection diffusion problem

Convection diffusion equations model the transport of a scalar quantity - denoted by u - which may be a concentration of a species or the temperature. It is a conservation law which accounts for two superimposed mechanisms: The first one, the convection, is the transport of the scalar quantity along a given stream $\underline{b}$ and is a directed phenomena (information moves downwind, information speed is limited). The corresponding convective flux is simply $\underline{b}u$. Convection is superimposed by diffusion which describes the exchange of momentum, concentration or temperature of neighbouring particles on a molecular level and so is not directed (information goes in all directions and propagates infinitely fast). The corresponding diffusive flux is $-\varepsilon\underline{\nabla}u$, where ε is a scalar for isotropic diffusion, which is the standard case. The right hand side of the convection diffusion equation describes sources and sinks for the scalar quantity. So if you balance the change of the scalar in time and the total in- and outgoing fluxes with sources and sinks on an arbitrary control volume you end up with the convection diffusion equation:

$$\frac{\partial u}{\partial t} + \mathsf{div}(-\varepsilon\underline{\nabla}u + \underline{b}u) = f$$

This type of equation arises in many flow problems and other physical applications like semiconductors. The steady Navier-Stokes equations which will be discussed later in chapter 2 are a vector-valued version of the convection diffusion problem, where the momentum ρu is the quantity which is transported with the velocity u and superimposed by viscous effects which are of diffusive character.

Our objective in this part is to solve the following scalar convection diffusion equation (with appropriate boundary and initial conditions) numerically. The problem reads

$$\begin{cases} \frac{\partial u}{\partial t} + \mathsf{div}(-\varepsilon\underline{\nabla}u + \underline{b}u) & = f & \text{in } \Omega \times [t_0, T], \quad \mathsf{div}(\underline{b}) = 0 \\ u & = u_D & \text{on } \Gamma_D \times [t_0, T] \\ \frac{\partial u}{\partial\underline{n}} & = g & \text{on } \Gamma_N \times [t_0, T] \\ \frac{\partial u}{\partial\underline{n}} + \beta u & = h & \text{on } \Gamma_R \times [t_0, T] \\ u & = u_0 & \text{on } \Omega \times t_0 \end{cases} \qquad (1.1.1)$$

where Ω is a bounded open domain in $\mathbb{R}^d$, $d = 1, 2, 3$ with boundary $\partial\Omega = \Gamma_D \cup \Gamma_N \cup \Gamma_R$. The boundary parts Γ_D, Γ_N and Γ_R denote the part of the boundary where Dirichlet, Neumann and Robin boundary conditions are prescribed respectively. $[t_0, T] \subset \mathbb{R}$ describes the time interval, where t_0 is the start time where the initial condition u_0 is prescribed and T is some later time.

The spatial domain is decomposed into a shape regular partition $\mathcal{T}_h$ of Ω consisting of simplices T. The element interfaces and element boundaries coinciding with the domain boundary are called *facets*. Furthermore the set of those facets E is denoted by $\mathcal{F}_h$ and there holds $\bigcup_{T \in \mathcal{T}_h} \partial T = \bigcup_{E \in \mathcal{F}_h} E$.

Note that we always assume that the convective velocity $\underline{b}$ is divergence-free which is reasonable if we assume that it comes from an incompressible flow field.

In this chapter we will only consider steady problems, i.e. $\frac{\partial u}{\partial t} = 0$ for now. The structure of this chapter is as follows:

First we will look at one limit of the convection diffusion equation, the poisson's equation, where the convective velocity is set to zero ($\|b\| = 0$) and only the purely diffusive phenomenon is described. The Hybrid DG formulation is introduced, discussed and a complete a priori error analysis is carried out. The other limit of the convection diffusion equation ($\varepsilon = 0$), which is the pure convective (linear and hyperbolic) one, will be the topic of the next section where the discretization of the convection with hybrid DG is presented and analyzed. Both cases will be joined afterwards to handle the full convection diffusion equation and the a priori error analysis for this case will be reconstructed from joining the analysis of both limits. As the major advantages and key properties of the Hybrid DG method are the same in all three parts, at least from a computational point of view, computational aspects will be discussed in more detail after the a priori error analysis. We conclude each section with numerical examples which shall highlight the features of the proposed Hybrid DG method.

1.1.1 State of the Art

For most elliptic problem the standard Finite Element method which we will call CG for "continuous Galerkin" is a convenient choice. Nevertheless there are situations where discontinuous formulations have their advantages. The CG method lacks desireable properties like conservation and the ability to represent "nearly discontinuities" on coarse meshes which might come up at material interfaces for example. An additional problem appears as soon as you leave the pure elliptic problems to convection diffusion problems. In the pure convective limit Discontinuous Galerkin methods which we call DG methods seem to be the more convenient choice, then. So you probably want to use a methods which can naturally be extended by DG formulations. To circumvent problems of the standard Finite Element method with convection diffusion equations there are conforming approaches like *streamline upwind Petrov Galerkin*, *Galerkin least squares* and *adjoint Galerkin least squares* methods (see [ESW05],[DH03]) which base upon the idea of adding diffusion numerically only in streamline direction and only in a consistent manner by formally exchanging the test functions of the variational formulation. Those methods certainly solve the stability problems of CG, but are often criticized to be overdiffusive.

A more general approach is the use of Standard DG formulations for elliptic operators like they are discussed in [ABCM02]. As hyperbolic equations can be included more naturally to DG methods than to conforming methods, the way to convection diffusion equations for DG methods is easy. Nevertheless this convenience comes with the prize of more unknowns and larger stencils.

Another approach was proposed by Egger and Schöberl in [ES09] where a mixed method for the diffusive part in hybridized form is coupled with a hybridized DG upwind method for the convective part. Here, even more unknowns have to be considered as in the DG case, but the stencil and so the nonzero entries in the matrices are dramatically reduced.

Our approach is strongly connected to the last two approaches. We use a Hybrid Discontinuous Galerkin method for the elliptic problem like they are presented and discussed in [CGL08] and the same hybridized DG upwind method as in [ES09] and couple both.

Another also promising idea is presented in [DG10]. Therein a Discontinuous Galerkin method with test functions differing from the trial functions is used, where the test functions are specifically tailored for stability resulting in a Discontinuous Petrov-Galerkin formulation which achieves L^2 best approximation for the linear transport equation. Thus the method also convergences optimally in the L^2 norm which does not hold for the above mentioned methods (including ours).

1.2 The Hybrid Discontinuous Galerkin method for second order elliptic problems

1.2.1 Introducing the method

In this section we consider poisson's equation:

$$\begin{cases} -\mathsf{div}(\varepsilon \underline{\nabla} u) & = f & \text{in } \Omega \\ u & = u_D & \text{on } \Gamma_D \\ \frac{\partial u}{\partial \underline{n}} & = g & \text{on } \Gamma_N \\ \frac{\partial u}{\partial \underline{n}} + \beta u & = h & \text{on } \Gamma_R \end{cases} \tag{1.2.1}$$

As we want to use a discontinuous Finite Element approximation of the weak solution $u \in H^1(\Omega)$, we look at the space of (element-)piecewise $H^1(T)$ functions which form the broken Sobolev space

$$H^1(\mathcal{T}_h) := \{u \in L^2(\Omega), u \in H^1(T) \ \forall T \in \mathcal{T}_h\}$$

On each element we approximate the $H^1(T)$ functions u with functions u_h which are polynomials up to degree $k \in \mathbb{N}$, but discontinuous. Thus there only holds

$$u_h|_T \in \mathcal{P}^k(T) \ \forall \ T \in \mathcal{T}_h \quad \Rightarrow \quad u_h \in H^1(\mathcal{T}_h) \subset L^2(\Omega) \tag{1.2.2}$$

When solving poisson's equation with CG you multiply the equation by a test function and integrate by parts on the whole domain to obtain a weak formulation. Due to the fact that the functions are no longer assumed to be $H^1(\Omega)$ regular but only $H^1(\mathcal{T}_h)$ regular, partial integration on the whole domain can no longer be applied. Thus, we integrate by parts on each element $T \in \mathcal{T}_h$, instead.

Starting with the problem

$$\text{Find } u \in H^1(\Omega) \text{ s.t. } -\int_\Omega \mathsf{div}(\varepsilon \underline{\nabla} u) \ v \, d\underline{x} = \int_\Omega f \ v \, d\underline{x} \ \forall v \in H_0^1(\Omega)$$
$$\text{with } H_0^1(\Omega) := \{u \in H^1(\Omega), u = 0 \text{ on } \Gamma_D\}$$

and formally doing integration by parts on each element $T \in \mathcal{T}_h$ we get a new variational problem for $u \in H^1(\Omega)$ and $v \in H^1(\mathcal{T}_h)$

$$-\sum_{T \in \mathcal{T}_h} \int_T \mathsf{div}(\varepsilon \underline{\nabla} u) \ v \, d\underline{x} = \sum_{T \in \mathcal{T}_h} \left\{ \int_T \varepsilon \ \underline{\nabla} u \ \underline{\nabla} v \, d\underline{x} - \int_{\partial T} \varepsilon \frac{\partial u}{\partial \underline{n}} \ v \, d\underline{s} \right\} = \sum_{T \in \mathcal{T}_h} \int_T f \ v \, d\underline{x} \tag{1.2.3}$$

(a) linear 2D CG DOF   (b) linear 2D DG DOF   (c) linear 2D HDG DOF
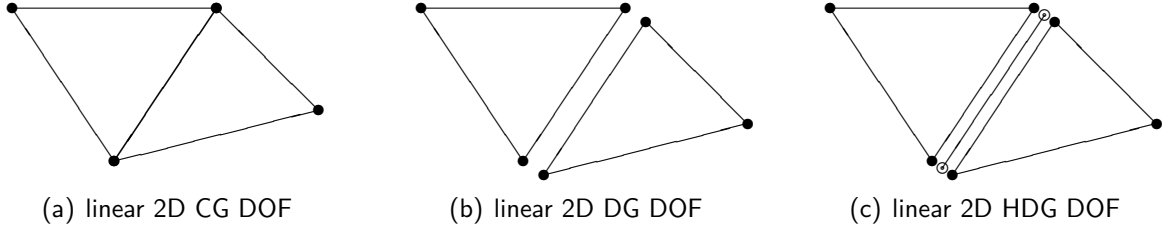
Figure 1.2.1: Degrees of freedom for different schemes

Note that we don't move the integral terms of the domain boundary on the right hand side. Thus we don't get Neumann boundary conditions as natural boundary conditions yet.

At this point we introduce additional unknown functions $u_F \in L^2(\mathcal{F}_h)$. Those functions only exist on the facets and for the true solution they represent the trace of u. The computational advantages of those additional *facet* functions, which are the essential ingredients of Hybrid DG, will be discussed later. In order to illustrate the situation, the degrees of freedom(DOF) of this method compared to those of CG and DG (for linear approximations in 2D) are schematically drawn in figure 1.2.1.

The above formulation obviously requires more regularity than $u \in H^1(\mathcal{T}_h)$, namely $u \in H^s(\mathcal{T}_h)$ with $s \geq \frac{3}{2}$, in order to make the normal derivative well defined. But, for ease of presentation, we will assume $u \in H^2(\mathcal{T}_h)$.

The assumption $u \in H^1(\Omega)$ is now extended by an additional $H^2(\mathcal{T}_h)$ regularity assumption. The spaces we will use on the continuous and the discrete level are V and V_h and are defined in the following.

$$
\begin{aligned}
V &:= \left\{ (u, u_F) : u \in H^2(\mathcal{T}_h) \cap H^1(\Omega), u_F \in L^2(\mathcal{F}_h) \right\} \\
V_h &:= \left\{ (u, u_F) : u \in \mathcal{P}^k(T) \; \forall \; T \in \mathcal{T}_h, u_F \in \mathcal{P}^k(E) \; \forall \; E \in \mathcal{F}_h \right\}
\end{aligned}
\tag{1.2.4}
$$

Here, $u \in H^2(\mathcal{T}_h)$ implies that u is in $H^2(T)$ for every element T of $\mathcal{T}_h$, but no continuity constraints are imposed on element interfaces. Dirichlet boundary conditions are posed on the facet functions only. Then, the spaces with (inhomogeneous and homogeneous) Dirichlet boundary conditions for the facet unknowns in it[1] can be defined analogously:

$$
\begin{aligned}
V_D &:= \{(u, u_F) \in V, u_F = u_D \text{ on } \Gamma_D\} & V_0 &:= \{(u, u_F) \in V, u_F = 0 \text{ on } \Gamma_D\} \\
V_{h,D} &:= \{(u, u_F) \in V_h, u_F = u_D \text{ on } \Gamma_D\} & V_{h,0} &:= \{(u, u_F) \in V_h, u_F = 0 \text{ on } \Gamma_D\}
\end{aligned}
\tag{1.2.5}
$$

With these choices for the discrete spaces another problem with the current form of the new formulation becomes evident: functions of neighbouring elements are independent from each other, such that all (element-)piecewise constant functions lie in the kernel of the left hand side bilinearform. So, we have to put some additional work into the formulation to overcome this problem. In essence, we achieve that by adding some consistent terms to the bilinearform. They either make use of the fact, that the exact solution is continuous across element interfaces or that the flux $-\varepsilon \nabla u$ is normal-continuous.

---

[1]imposing boundary conditions strong for the facet functions is like imposing them weakly in a Standard DG manner (see section 1.2.2.2)

The latter justifies the following relation:

$$\sum_{T \in \mathcal{T}_h} \int_{\partial T} \varepsilon \frac{\partial u}{\partial \underline{n}} \, v_F \, d\underline{s} = \underbrace{\sum_{E \in \mathcal{F}_h^{int}} \left\{ \int_E \varepsilon \frac{\partial u}{\partial \underline{n}^+} \, v_F \, d\underline{s} + \int_E \varepsilon \frac{\partial u}{\partial \underline{n}^-} \, v_F \, d\underline{s} \right\}}_{=0 \text{ for true solution}}$$

$$+ \sum_{E \in \mathcal{F}_h^{ext}} \int_E \varepsilon \frac{\partial u}{\partial \underline{n}} \, v_F \, d\underline{s} = \int_{\partial \Omega} \varepsilon \frac{\partial u}{\partial \underline{n}} \, v_F \, d\underline{s}$$

(1.2.6)

where we used $\mathcal{F}^{ext} := \{ E \in \mathcal{F}_h; \ E \cap \partial \Omega \neq \emptyset \}$ and $\mathcal{F}^{int} := \{ E \in \mathcal{F}_h; \ E \cap \partial \Omega = \emptyset \}$ for the set of all interior and exterior facets.

For inner facets we get element boundary integrals from both sides which cancel out for the true solution. Thus if we add (1.2.6) to (1.2.3) and plug in the boundary conditions for the new integral terms and introduce $[\![\mathbf{v}]\!] := v - v_F$ where the abbreviation $\mathbf{v}$ in contrast to the expression (v, v_F) is written bold:

Find $(u, u_F) \in V_D$, such that for all $(v, v_F) \in V_0$ there holds:

$$\sum_{T \in \mathcal{T}_h} \left\{ \int_T \varepsilon \, \underline{\nabla} u \, \underline{\nabla} v \, d\underline{x} - \int_{\partial T} \varepsilon \frac{\partial u}{\partial \underline{n}} \, [\![\mathbf{v}]\!] \, d\underline{s} \right\} + \beta \int_{\Gamma_R} \varepsilon \, u \, v_F \, d\underline{s}$$

(1.2.7)

$$= \sum_{T \in \mathcal{T}_h} \int_T f \, v \, d\underline{x} + \int_{\Gamma_N} \varepsilon \, g \, v_F \, d\underline{s} + \int_{\Gamma_R} \varepsilon \, h \, v_F \, d\underline{s}$$

Due to (1.2.6), Neumann boundaries appear as natural boundary conditions in this variational formulation again. We add additional terms, which now make use of the $H^1(\Omega)$-regularity of the exact solution, which ensures that for $(u, u_F) \in V$ there holds $[\![\mathbf{u}]\!] = u - u_F = u - tr_{\partial T}(u) = 0$ on each facet. Thus by adding integrals on each facet, which contain this term, we add only consistent terms, and then element neighbours couple, even though only indireclty, with each other. We choose to add

$$- \int_{\partial T} \varepsilon \frac{\partial v}{\partial \underline{n}} \, [\![\mathbf{u}]\!] \, d\underline{s} \qquad \text{for symmetry}$$

and

$$\int_{\partial T} \tau_h \varepsilon \, [\![\mathbf{u}]\!] \, [\![\mathbf{v}]\!] \, d\underline{s} \qquad \text{for stability}$$

with some stabilization parameter τ_h, which has to be chosen appropriately.

We can also make use of the consistency relation to exchange u with u_F at the Robin boundary integral. After those manipulations we get the bilinearform

$$\mathcal{B}_h((u, u_F), (v, v_F)) := \sum_{T \in \mathcal{T}_h} \left\{ \int_T \varepsilon \, \underline{\nabla} u \, \underline{\nabla} v \, d\underline{x} - \int_{\partial T} \varepsilon \frac{\partial u}{\partial \underline{n}} \, [\![\mathbf{v}]\!] \, d\underline{s} \right.$$

(1.2.8)

$$\left. - \int_{\partial T} \varepsilon \frac{\partial v}{\partial \underline{n}} \, [\![\mathbf{u}]\!] \, d\underline{s} + \int_{\partial T} \varepsilon \, \tau_h \, [\![\mathbf{u}]\!] \, [\![\mathbf{v}]\!] \, d\underline{s} \right\} + \beta \int_{\Gamma_R} \varepsilon \, u_F \, v_F \, d\underline{s}$$

The insertion of the last two consistent integrals are degrees of freedom to tune and control the method and so the proposed bilinearform is certainly not the only reasonable choice you could make. In the section 1.2.2 we will try to understand the properties and the relations of this formulation to other methods. As this method has some obvious similarities to the *Standard DG* symmetric interior penalty method, we also want to mention that some other DG methods like the NIPG method or the DG method by Baumann and Oden (see also [ABCM02]) can be easily translated to a similar hybridized

version.

The most important properties of the proposed formulation are symmetry, coercivity for sufficiently large τ_h (we will discuss this later) and consistency. Symmetry is obvious, consistency and coercivity will be further discussed in section 1.2.3.1 and 1.2.3.2.

For ease of notation, we will often denote the pair (u, u_F) as $\mathbf{u}$ and also shorten the notation for the bilinearform by $\mathcal{B}_h(\mathbf{u}, \mathbf{v})$. Thus our discrete problem reads

Find $\mathbf{u} \in V_{h,D}$, such that for all $\mathbf{v} \in V_{h,0}$:

$$\mathcal{B}_h(\mathbf{u}, \mathbf{v}) = \langle \tilde{f}, v \rangle = \sum_{T \in \mathcal{T}_h} \int_T f\, v\, d\underline{x} + \int_{\Gamma_N} \varepsilon\, g\, v_F\, d\underline{s} + \int_{\Gamma_R} \varepsilon\, h\, v_F\, d\underline{s} \tag{1.2.9}$$

Several advantages come up when using this kind of Hybrid DG methods:

- Conservation of a discrete flux (see 1.2.4) like Standard DG or mixed methods

- Standard FEM like elementwise matrix assembling as neighbouring degrees of freedom don't couple directly (see 1.5.2). Inner degrees of freedom can also be eliminated by static condensation.

- Extensions to DG schemes for convection diffusion problems can be considered for quite easily

Remark 1.2.1 (Stabilization paramater):
For a sufficiently large stabilization parameter $\tau_h > \tau_h^*$ the method is stable. This parameter depends on the size and shape of the neighbouring elements of a facet. From scaling arguments it is also clear that τ_h should scale with $\frac{1}{h}$ where h is a characteristic length of the neighbouring elements. When we show coercivity in 1.2.3.2 we'll see that the τ_h^* can be explicitly derived from inverse inequalities.

Remark 1.2.2 (Non-Uniqueness of the Abbreviation "HDG method"):
In this work the proposed symmetric interior penalty HDG method will be adressed as the "HDG method" even though the correct designation would be the somewhat more lenghty expression "hybridized symmetric interior penalty Discontinuous Galerkin method", which distinguishes it from other HDG methods that could also be used (see [CGL08]) . Most of the important results presented in this work that hold for our particular choice, the hybridized symmetric interior penalty DG method, can be translated to other HDG methods as well.

Remark 1.2.3 (Dirichlet boundary conditions):
If Dirichlet boundary conditions are used corresponding degrees of freedom can be removed from the formulation. This can be done by shifting domain boundary integrals of the bilinearform $\mathcal{B}_h$ on the r.h.s.. The necessary adjustments on the linearform $\langle \tilde{f}, v \rangle$ then lead to a modified variational problem, which incorporates Dirichlet boundary conditions in a "Nitsche" way. Another possibility is to plug in the boundary conditions on the matrix-vector level after discretizing the variational formulation . Of course, both ways end up with the same solution. However, our analysis will be carried out including Dirichlet boundary facet functions, thus the bilinearform does not have to distinguish between exterior and interior facets, which simplifies the notation.

1.2.2   Characterizations of the method

The HDG method is deeply connected to other Finite Element methods and can be derived in different ways. We want to briefly highlight the relation to (hybridized) mixed methods, Standard DG methods and an alternative HDG formulation. As we will see, the proposed method is actually equal to one specific hybridized mixed method, one specific Standard DG method and also to the alternative HDG method. Nevertheless the methods are derived from different point of views. At the end of the section, after the several relations to other Finite Element methods are discussed, an attemp to visualize the interconnection between the different formulations, and thereby summarize this section, is made.

For ease of presentation we don't draw special attention on the treatment of boundary conditions, meaning we assume homogeneous Dirichlet boundary conditions on the whole boundary. However, for all three approaches different boundary conditions can be easily considered for.

1.2.2.1   HDG as a modified hybridized mixed method

Mixed methods
As a starting point we reformulate the second order problem $-\text{div}(\varepsilon\underline{\nabla}u) = f$, equation (1.2.1) as a first order system by substituting the flux $\underline{\sigma} = -\varepsilon\underline{\nabla}u$

$$\begin{cases} \dfrac{1}{\varepsilon}\underline{\sigma} & + & \underline{\nabla}u & = & 0 \\ \text{div}(\underline{\sigma}) & & & = & f \end{cases} + b.c. \tag{1.2.10}$$

The appropriate spaces for a weak solution are $H(\text{div},\Omega)$ for $\underline{\sigma}$ and $L^2(\Omega)$ for u. After test with $\underline{\tau}$ and v) and applying partial integration on the gradient part, we end up with the standard mixed method for the poisson equation[2,3]

$$\begin{cases} (\frac{1}{\varepsilon}\underline{\sigma},\underline{\tau})_\Omega & - & (u,\text{div}(\underline{\tau}))_\Omega & = & (-u_D,\tau_n)_{\partial\Omega} & \forall\ \underline{\tau}\in H(\text{div},\Omega) \\ -(\text{div}(\underline{\sigma}),v)_\Omega & & & = & (-f,v)_\Omega & \forall\ v\in L^2(\Omega) \end{cases} \tag{1.2.11}$$

The appropriate discrete spaces Σ_h and Q_h are $H(\text{div},\Omega)$-conforming and discontinous Finite Element spaces, respectively.

$$\Sigma_h \subset H(\text{div},\Omega)\ ,\qquad Q_h \subset L^2(\Omega)$$

If we choose Σ_h to be an $H(\text{div},\Omega)$-conforming finite element space of order[4] $k+1$, which we'll label with Σ_h^{k+1} and Q_h to be the piecewise polynomial space up to order k, which we'll label with Q_h^k, you can show stability and optimal convergence rates for the error $(\|u-u_h\|_{L^2}^2 + \|\sigma-\sigma_h\|_{L^2}^2)^{\frac{1}{2}}$ (see standard text books, e.g. [Brae97]). The construction of $H(\text{div},\Omega)$-conforming Finite Element spaces will be discussed in more detail in section 2.2.

One advantage of this method is the conservation of the flux σ. But the disadvantage compared to CG is that you have solve a saddle point problem with considerably more degrees of freedom. This disadvantage can be circumvented with hybridized mixed methods, which we'll discuss next.

---

[2]for standard mixed methods you often choose $\underline{\sigma} = \varepsilon\underline{\nabla}u$ and so you end up with different signs

[3]when looking at mixed systems we will use the inner product notation instead of integrals

[4]The order of an $H(\text{div},\Omega)$-conforming space describes the <u>maximum</u> polynomial order of the polynomials itself not of the divergence. RT$_0$ and BDM$_1$ are, by this definition, examples for Σ_h^1.

Hybridized mixed methods

For mixed methods the primal unknown is chosen to be discontinuous and the flux is introduced as a new unknown with normal continuity incorporated in the space Σ_h, so we directly get conservation of the mixed method. When using hybridization techniques, you break up the normal-continuity constraint for the Finite Element space and use lagrange multipliers to reimpose the according constraint. We achieve that by multiplying (1.2.10) with testfunctions and integrating elementwise over the whole domain. When doing integration by parts on every element, boundary terms for all element interfaces appear. We then define the trace of u on the boundary of each element as new unknowns u_F and use them as lagrange multipliers for the conservation-constraint $[\![\underline{\sigma}]\!]_n := \underline{\sigma}^+\underline{n}^+ + \underline{\sigma}^-\underline{n}^- = 0$ (in a weak sense):

Find $(\sigma, u, u_F) \in \Sigma_{h,disc}^{k+1} \times Q_h^k \times F_h^k$, such that

$$
\left\{
\begin{array}{llll}
\sum_{T \in \mathcal{T}_h} & (\frac{1}{\varepsilon}\underline{\sigma}, \underline{\tau})_T & -(u, \mathrm{div}(\underline{\tau}))_T \quad +(u_F, \tau_n)_{\partial T} & = & 0 & \forall \underline{\tau} \in \Sigma_{h,disc}^{k+1} \\
\sum_{T \in \mathcal{T}_h} & -(\mathrm{div}(\underline{\sigma}), v)_T & & = & (-f, v)_\Omega & \forall v \in Q_h^k \\
\sum_{T \in \mathcal{T}_h} & (\underline{\sigma} \cdot \underline{n}, v_F)_{\partial T} & & = & 0 & \forall v_F \in F_h^k
\end{array}
\right. \quad (1.2.12)
$$

where $F_h^k := \{v \in L^2(\mathcal{F}_h); v|_{\partial T} \in \mathcal{P}^k\}$ is the space of (facet-) piecewise polynomials. All element unknowns u and σ only couple with other unknowns of the same element or the facet unknowns u_F. The (element-)local problems are furthermore uniquely solvable (see [BF91]) and so the linear system can be reduced to the facet unknowns by static condensation. A remarkable property of the final linear system for the facet unknowns is the symmetric positive definiteness.

Until this point, you might see the hybridization as an implemenatation trick. But after solving the linear system for the facet unknowns, you can enhance your discrete solution even more: The solution u_F, which describes Dirichlet data to local problems, can be used for (element-)local postprocessings like they are discussed in [BF91] to gain an order $k + 1$ reconstruction of the primal variable u.

Hybrid DG methods

The third equation of (1.2.12) ensures continuity of $\underline{\sigma}$ and so conservation of the hybrid mixed method. Another possibility of ensuring conservation without enforcing continuity of $\underline{\sigma}$ is to pose a condition on a *numerical flux* $\hat{\underline{\sigma}}$ instead. To get there, we integrate both div-terms by parts and replace $\underline{\sigma} \cdot \underline{n}$ by $\hat{\underline{\sigma}} \cdot \underline{n}$. Now as we did integration by parts we have to adapt our Finite Element spaces. Now we choose the primal variable u of one order higher than $\underline{\sigma}$, s.t. $\underline{\nabla} u$ and $\underline{\sigma}$ lie in the same space. The new problem reads:

Find $(\underline{\sigma}, u, u_F) \in [Q_h^{k-1}]^d \times Q_h^k \times F_h^k$, such that

$$
\left\{
\begin{array}{llll}
\sum_{T \in \mathcal{T}_h} & (\frac{1}{\varepsilon}\underline{\sigma}, \underline{\tau})_T & + (\underline{\nabla} u, \underline{\tau})_T \quad + (u_F - u, \tau_n)_{\partial T} & = & 0 & \forall \underline{\tau} \in [Q_h^{k-1}]^d \\
\sum_{T \in \mathcal{T}_h} & (\underline{\sigma}, \underline{\nabla} v)_T & - (\hat{\underline{\sigma}} \cdot \underline{n}, v)_{\partial T} & = & (-f, v)_\Omega & \forall v \in Q_h^k \\
\sum_{T \in \mathcal{T}_h} & (\hat{\underline{\sigma}} \cdot \underline{n}, v_F)_{\partial T} & & = & 0 & \forall v_F \in F_h^k
\end{array}
\right. \quad (1.2.13)
$$

Here you can use several appropriate fluxes. Choosing $\hat{\underline{\sigma}} = \underline{\sigma}$ would lead to a formulation which is equivalent to standard mixed methods again. If we test our primal HDG-IP formulation (1.2.9) and (1.2.8) with $(0, v_F)$, we see that we already (implicitly) used the third equation of (1.2.13), s.t. our numerical flux is

$$
\hat{\underline{\sigma}} \cdot \underline{n} = -\varepsilon \frac{\partial u}{\partial \underline{n}} + \varepsilon \tau_h (u - u_F) \cdot \underline{n} \qquad (1.2.14)
$$

With the linear and element-local lifting operator $\mathcal{L} : L^2(\partial T) \to H(\mathrm{div}, T)$ which fulfills

$$
\int_T \mathcal{L}(w)\, \underline{\tau}\, d\underline{x} = \int_{\partial T} w\, \underline{\tau} \cdot \underline{n}\, d\underline{s} \qquad \forall \underline{\tau} \in H(\mathrm{div}, T) \qquad (1.2.15)
$$

we can formally rewrite the first equation of (1.2.13) as

$$\int_T \left(\frac{1}{\varepsilon}\underline{\sigma} + \underline{\nabla}u\right)\underline{\tau}\, d\underline{x} + \int_{\partial T}(u_F - u)\,\underline{\tau}\cdot\underline{n}\,d\underline{s} = \int_T \left(\frac{1}{\varepsilon}\underline{\sigma} + \underline{\nabla}u + \underline{\mathcal{L}}(u_F - u)\right)\underline{\tau}\, d\underline{x} = 0 \quad \forall\,\underline{\tau}\in[Q_h^{k-1}]^d$$

(1.2.16)

Thus the strong form for $\underline{\sigma}$ reads:

$$\underline{\sigma} = -\varepsilon\underline{\nabla}u + \varepsilon\underline{\mathcal{L}}(u - u_F) \tag{1.2.17}$$

Plugging (1.2.14) and (1.2.17) into the second equation of (1.2.13) we get $\forall v\in Q_h^k$:

$$\sum_{T\in\mathcal{T}_h}\left\{\int_T \varepsilon\underline{\nabla}u\,\underline{\nabla}v\, d\underline{x} - \int_T \varepsilon\underline{\mathcal{L}}(u - u_F)\,\underline{\nabla}v\, d\underline{x} - \int_{\partial T}\varepsilon\frac{\partial u}{\partial\underline{n}}\,v\,d\underline{s} + \int_{\partial T}\tau_h(u - u_F)\,v\,d\underline{s}\right\} = \int_\Omega f\,v\,d\underline{x}$$

(1.2.18)

If we plug the numerical flux of (1.2.14) also into the third equation of (1.2.13) and multiply by -1 we get:

$$\sum_{T\in\mathcal{T}_h} -\int_T \hat{\underline{\sigma}}\cdot\underline{n}\,v_F\,d\underline{x} = \sum_{T\in\mathcal{T}_h}\left\{\int_{\partial T}\varepsilon\frac{\partial u}{\partial\underline{n}}\,v_F\,d\underline{s} - \int_{\partial T}\tau_h(u - u_F)\,v_F\,d\underline{s}\right\} = 0 \tag{1.2.19}$$

Adding equation (1.2.18) and (1.2.19) together, we end up with the primal formulation of Hybrid DG again (compare (1.2.8))

$$\mathcal{B}_h(\mathbf{u},\mathbf{v}) := \sum_{T\in\mathcal{T}_h}\left\{\int_T \varepsilon\,\underline{\nabla}u\,\underline{\nabla}v\, d\underline{x} - \int_{\partial T}\varepsilon\,\frac{\partial u}{\partial\underline{n}}\,[\![\mathbf{v}]\!]\,d\underline{s} - \int_{\partial T}\varepsilon\,\frac{\partial v}{\partial\underline{n}}\,[\![\mathbf{u}]\!]\,d\underline{s} + \int_{\partial T}\varepsilon\tau_h\,[\![\mathbf{u}]\!]\,[\![\mathbf{v}]\!]\,d\underline{s}\right\}$$

(1.2.20)

where $\mathbf{u}$ and $\mathbf{v}$ are taken out of $Q_h^k \times F_h^k = V_h^k$

Remark 1.2.4 (Reducing the number of facet unknowns):
If we compare our approximation abilities with respect to the facet unknowns' polynomial degree, we notice that for the hybridized mixed method we can achieve approximation of order $k + 1$ for u (after postprocessing) when using polynomials of order k on the facets. In (1.2.13) we just achieve approximation of order k when using polynomials of order k on the facets, so we want to briefly discuss how to overcome this drawback.
The reason we take u_F and v_F of polynomial degree k is that $\hat{\underline{\sigma}}\cdot\underline{n}$ is of degree k(see third equation of (1.2.13)). If we modify $\hat{\underline{\sigma}}\cdot\underline{n}$ s.t. it is only of polynomial degree $k - 1$, the test function v_F and so the lagrange multiplier u_F can be taken from F_h^{k-1}, what is our goal here. So the global degrees of freedom (meaning those which can not be condensated out on the element level) could be reduced further. To achieve this, we could introduce an L^2-orthogonal projector P_{k-1}, which projects into the space of (facet-) piecewise polynomials of order $k - 1$ and adjust our numerical flux:

$$\hat{\underline{\sigma}}\cdot\underline{n} = -\varepsilon\frac{\partial u}{\partial\underline{n}} + \varepsilon\tau_h P_{k-1}(u_h - u_F) \tag{1.2.21}$$

With this flux and the L^2-orthogonality of the projector we get a new bilinearform for which we now take $\mathbf{u}$ and $\mathbf{v}$ in $Q_h^k \times F_h^{k-1}$:

$$\mathcal{B}_h(\mathbf{u},\mathbf{v}) := \sum_{T\in\mathcal{T}_h}\left\{\int_T \varepsilon\,\underline{\nabla}u\,\underline{\nabla}v\, d\underline{x} - \int_{\partial T}\varepsilon\,\frac{\partial u}{\partial\underline{n}}\,[\![\mathbf{v}]\!]\,d\underline{s} - \int_{\partial T}\varepsilon\,\frac{\partial v}{\partial\underline{n}}\,[\![\mathbf{u}]\!]\,d\underline{s} + \int_{\partial T}\tau_h\varepsilon\,P_{k-1}[\![\mathbf{u}]\!]\,P_{k-1}[\![\mathbf{v}]\!]\,d\underline{s}\right\}$$

This "optimization" of the bilinearform only makes sense as long as we have pure elliptic problems. As soon as we add convection we need u_F and v_F to be of order k again. That's why the a priori error analysis does not include this "optimization".

1.2.2.2 HDG as a Standard DG method

For standard Discontinuous Galerkin methods, there is a lot of work going on for many years and a lot of theory has been derived for them. In this small part we want to derive a representation of the proposed HDG method which fits into the framework of Standard DG methods. As we presented our HDG method in a primal formulation, the reformulation as a Standard DG formulation in primal form is the most natural approach. From the characterization of the HDG method as a hybridized mixed method, we know that the equations corresponding to the facet degrees of freedom ensure conservation of a numerical flux. But you can also derive an expression for the lagrange multipliers u_F from those equations. By doing so we eliminate u_F, return to Standard DG unknowns and get a primal DG formulation. Afterwards, we will conclude this excursus with the numerical trace $\hat{u}$ and flux $\hat{\sigma}$ of the HDG method.

Eliminating the facet unknowns u_F
We eliminate the facet unknowns u_F of a discrete solution of the variational problem. We can do that for a facet $E \in \mathcal{F}_h$ by testing equation (1.2.9), (1.2.8) with $\mathbf{v} = (0, v_F)$ where $v_F \in L^2(E)$. Adding up the terms of two neighbouring elements we get (with $\{\!\{\cdot\}\!\}$ the arithmetic average of both element values)

$$\int_E [\![\varepsilon \underline{\nabla} u]\!] \, v_F \, d\underline{s} - 2 \int_E \left(\{\!\{\tau_h \varepsilon u\}\!\} - \{\!\{\tau_h \varepsilon\}\!\} u_F \right) \, v_F \, d\underline{s} = 0 \tag{1.2.22}$$

In strong form [5] this means

$$u_F = \frac{\{\!\{\tau_h \varepsilon u\}\!\}}{\{\!\{\tau_h \varepsilon\}\!\}} - \frac{1}{2} \frac{[\![\varepsilon \underline{\nabla} u]\!]}{\{\!\{\tau_h \varepsilon\}\!\}} \tag{1.2.23}$$

DG represenation of the HDG-jump operator $[\![\mathbf{u}]\!]$
Now we want to make use of this representation and plug it into the bilinearform so that we end up with a formulation which is directly comparable to other DG methods. For ease of presentation, we will assume $\varepsilon \equiv 1$ and $\tau_h|_E \equiv const$ on each facet E, s.t. $u_F = \{\!\{u\}\!\} - \frac{1}{2}[\![\underline{\nabla} u]\!]$. Plugging relation (1.2.23) into the bilinearform $\mathcal{B}_h(\cdot, \cdot)$, we get rid of the additional unknowns u_F and so the additional test functions on the facets are no longer required. All jump terms for the test functions $[\![\mathbf{v}]\!] = v - v_F$ now just get v and $[\![\mathbf{u}]\!]$ is:

$$[\![\mathbf{u}]\!] = u - u_F = \frac{1}{2} \left([\![u]\!]^* + \frac{1}{\tau_h} [\![\underline{\nabla} u]\!]_n \right)$$

where we used $u - \{\!\{u\}\!\} = \frac{1}{2}[\![u]\!]^*$ with $[\![u]\!]^* := u - u^{ext}$ the DG jump operator and u^{ext} the neighbour elements' value[6]. This gives us the *Standard DG* Bilinearform

$$\mathcal{B}_h^{DG}(u, v) := \sum_{T \in \mathcal{T}_h} \underbrace{\int_T \underline{\nabla} u \, \underline{\nabla} v \, d\underline{x}}_{A} - \underbrace{\int_{\partial T} \frac{\partial u}{\partial \underline{n}} v \, d\underline{s}}_{B} - \underbrace{\int_{\partial T} \frac{\partial v}{\partial \underline{n}} [\![\mathbf{u}]\!] \, d\underline{s}}_{C} + \underbrace{\int_{\partial T} \tau_h [\![\mathbf{u}]\!] v \, d\underline{s}}_{D} \tag{1.2.24}$$

Straight forward computation gives us the following relations for the sum of element boundary integrals and according facet integrals. Here the terms with $[\![\mathbf{u}]\!]$ are devided into two parts. One with $\frac{1}{2}[\![u]\!]^*$

---

[5] To come to the strong form we have to assume that the test space is sufficiently large. In our applications this is always true.

[6] in contrast to $[\![\underline{\nabla} u]\!]_n$ the sign of $[\![u]\!]^*$ depends on the element from where you look at the facet. That's what we indicate by the star.

which is marked as part 1 and the one with $\frac{1}{2\tau_h}[\![\nabla u]\!]_n$ is marked as part 2follows:

$$
\begin{aligned}
B: \quad &-\sum_{T\in\mathcal{T}_h}\int_{\partial T}\frac{\partial u}{\partial \underline{n}}v\,d\underline{s} &&= -\sum_{E\in\mathcal{F}_h}\int_E \{\!\{\nabla u\}\!\}^*[\![v]\!]^* + \{\!\{v\}\!\}[\![\nabla u]\!]_n\,d\underline{s}\\
C.1: \quad &-\frac{1}{2}\sum_{T\in\mathcal{T}_h}\int_{\partial T}[\![u]\!]^*\frac{\partial v}{\partial \underline{n}}\,d\underline{s} &&= -\sum_{E\in\mathcal{F}_h}\int_E [\![u]\!]^*\{\!\{\nabla v\}\!\}^*\,d\underline{s}\\
C.2: \quad &-\sum_{T\in\mathcal{T}_h}\int_{\partial T}\frac{1}{2\tau_h}\frac{\partial v}{\partial \underline{n}}[\![\nabla u]\!]_n\,d\underline{s} &&= -\sum_{E\in\mathcal{F}_h}\int_E \frac{1}{2\tau_h}[\![\nabla u]\!]_n[\![\nabla v]\!]_n\,d\underline{s}\\
D.1: \quad &\frac{1}{2}\sum_{T\in\mathcal{T}_h}\int_{\partial T}\tau_h[\![u]\!]^*v\,d\underline{s} &&= \sum_{E\in\mathcal{F}_h}\int_E \frac{\tau_h}{2}[\![u]\!]^*[\![v]\!]^*\,d\underline{s}\\
D.2: \quad &\frac{1}{2}\sum_{T\in\mathcal{T}_h}\int_{\partial T}[\![\nabla u]\!]_n v\,d\underline{s} &&= \sum_{E\in\mathcal{F}_h}\int_E [\![\nabla u]\!]_n\{\!\{v\}\!\}\,d\underline{s}
\end{aligned}
\tag{1.2.25}
$$

Primal DG formulation
Plugging all those relations into $\mathcal{B}_h^{DG}(u,v)$ gives the primal DG form

$$
\mathcal{B}_h^{DG}(u,v) = \sum_{T\in\mathcal{T}_h}\underbrace{\int_T \nabla u\,\nabla v\,d\underline{x}}_{A} + \sum_{E\in\mathcal{F}_h}\underbrace{-\int_F \{\!\{\nabla u\}\!\}^*\,[\![v]\!]^*\,d\underline{s}}_{B+D.2} \underbrace{-\int_F \{\!\{\nabla v\}\!\}^*\,[\![u]\!]^*\,d\underline{s}}_{C.1}
$$

$$
+ \underbrace{\frac{1}{2}\int_F \tau_h[\![u]\!]^*\,[\![v]\!]^*\,d\underline{s}}_{D.1} \underbrace{-\frac{1}{2}\int_F \frac{1}{\tau_h}[\![\nabla u]\!]^*\,[\![\nabla v]\!]^*\,d\underline{s}}_{C.2}
\tag{1.2.26}
$$

If we drop off the last integral and weight the stabilization parameter by a factor of two, we'd recover the *Standard DG (symmetric) Interior Penalty* formulation for the laplace operator. Thus we see that our proposed HDG (symmetric) Interior Penalty formulation is quiet similar but not identical.

Numerical flux and trace of the HDG method
In [ABCM02] the *Standard DG (symmetric) Interior Penalty* as well as other DG methods are compared by formulating a general Standard DG bilinearform with numerical traces and fluxes $\hat{u}$ and $\underline{\hat{\sigma}}$. If you replace u_F in equation (1.2.13) by a numerical trace $\hat{u}$ you have the general DG formulation in mixed form:

$$
\begin{cases}
\sum_{T\in\mathcal{T}_h} & (\frac{1}{\varepsilon}\underline{\sigma},\underline{\tau})_T & +(\nabla u,\underline{\tau})_T & +(\hat{u}-u,\tau_n)_{\partial T} & = & 0 & \forall \underline{\tau}\in[Q_h^{k-1}]^d\\
\sum_{T\in\mathcal{T}_h} & (\underline{\sigma},\nabla v)_T & -(\underline{\hat{\sigma}}\cdot\underline{n},v)_{\partial T} & & = & \sum_{T\in\mathcal{T}_h}(-f,v)_T & \forall v\in Q_h^k\\
\sum_{T\in\mathcal{T}_h} & (\underline{\hat{\sigma}}\cdot\underline{n},v_F)_{\partial T} & & & = & 0 & \forall v_F\in F_h^k
\end{cases}
\tag{1.2.27}
$$

For the different choices of those fluxes $\underline{\hat{\sigma}}$ and traces $\hat{u}$ you get different DG formulations. As was already pointed out in [CGL08, equation (3.27)] the proposed HDG formulation has the following numerical trace and flux:

$$
\hat{u} = u_F = \{\!\{u\}\!\} - \frac{1}{2\tau_h}[\![\nabla u]\!]_n \qquad \text{see (1.2.23) with } \varepsilon\equiv 1 \text{ and } \tau_h\equiv const \text{ on each facet}
$$

$$
\underline{\hat{\sigma}}\cdot\underline{n} = -\{\!\{\nabla u\}\!\}^* + \frac{\tau_h}{2}[\![u]\!]^* \qquad \text{average } \underline{\hat{\sigma}} \text{ from (1.2.14)}
\tag{1.2.28}
$$

Both values are consistent as they coincide with the trace of the solution and the trace of the normal derivative, respectively, for functions $u\in H^2(\Omega)$.

1.2.2.3   Alternative HDG formulation

As we have seen, the proposed representation of our HDG formulation can be regained from (1.2.27) by choosing a specific numerical flux and introducing $\hat{u} = u_F$ as a new unknown. You might also want to do it the other way around, set $\hat{\sigma} \cdot \underline{n} = \sigma_{n,F}$ and choose the numerical trace $\hat{u}$ as in (1.2.28). For now, we again want to assume $\varepsilon \equiv 1$ and $\tau_h|_E \equiv const$ on each facet E. With the new unknowns we also pose a new condition on the solution, which is [7]

$$\sum_{T \in \mathcal{T}_h} \left\{ \int_{\partial T} u \, \tau_{n,F} \, d\underline{s} - \int_{\partial T} \frac{1}{\tau_h} \left(\sigma_{n,F} + \frac{\partial u}{\partial n} \right) \tau_{n,F} \, d\underline{s} \right\} = 0 \tag{1.2.29}$$

With the relations

$$\{\!\{u\}\!\} = u - \frac{1}{2} [\![u]\!]^* \tag{1.2.30a}$$

$$\frac{1}{2} [\![\nabla u]\!]_n = \frac{\partial u}{\partial \underline{n}} - \{\!\{\underline{\nabla} u\}\!\}^* \tag{1.2.30b}$$

we can reformulate the numerical trace:

$$\begin{aligned}
\hat{u} &= \{\!\{u\}\!\} - \frac{1}{2\tau_h} [\![\nabla u]\!]_n \\
&\overset{(1.2.30)}{=} u - \frac{1}{\tau_h} \frac{\tau_h}{2} [\![u]\!]^* - \frac{1}{\tau_h} \left(\frac{\partial u}{\partial \underline{n}} - \{\!\{\underline{\nabla} u\}\!\}^* \right) \\
&= u - \frac{1}{\tau_h} \left(\frac{\partial u}{\partial \underline{n}} - \{\!\{\underline{\nabla} u\}\!\}^* + \frac{\tau_h}{2} [\![u]\!]^* \right) \\
&= u - \frac{1}{\tau_h} \left(\frac{\partial u}{\partial \underline{n}} + \sigma_{n,F} \right)
\end{aligned} \tag{1.2.31}$$

where we used $\sigma_{n,F} = \hat{\underline{\sigma}} \cdot \underline{n}$.

Now we can use the representation of the flux (1.2.17) when replacing u_F by $\hat{u}$. For $u - \hat{u}$ we get

$$[\![\mathbf{u}]\!] = u - u_F = u - \hat{u} = \frac{1}{\tau_h} \left(\frac{\partial u}{\partial \underline{n}} + \sigma_{n,F} \right) \tag{1.2.32}$$

And we can plug it into the representaion of the flux (1.2.17)

$$\underline{\sigma} = -\underline{\nabla} u + \frac{1}{\tau_h} \mathcal{L} \left(\frac{\partial u}{\partial \underline{n}} + \sigma_{n,F} \right) \tag{1.2.33}$$

Further putting this into the second equation of (1.2.27) and adding the new equation (1.2.29), we get for all $(v, \tau_{n,F}) \in Q_h^k \times \tilde{F}_h^k$:

$$\sum_T \left\{ \int_T \underline{\nabla} u \, \underline{\nabla} v \, d\underline{x} + \int_{\partial T} \sigma_{n,F} \, v \, d\underline{s} - \int_{\partial T} u \, \tau_{n,F} \, d\underline{s} + \int_{\partial T \tau_h} \frac{1}{\tau_h} \left(\sigma_{n,F} + \frac{\partial u}{\partial \underline{n}} \right) \left(\tau_{n,F} + \frac{\partial v}{\partial \underline{n}} \right) d\underline{s} \right\} = \int_\Omega f v \, d\underline{x} \tag{1.2.34}$$

Although this formulation is equivalent to our proposed method there are some computational disadvantages:

1. the problem cannot be reduced to $\sigma_{n,F}$ only, as the local problems are singular.

2. the system is indefinite

This approach was introduced in [EWY02] as a (stabilized) hybridized primal mixed method.

---

[7] we pose exactly this condition as we aim to get an equivalent formulation to the one in (1.2.8)

1.2.2.4 Summary of the Interconnection between the discussed Finite Element methods

After relations to some other Finite Element methods are discussed, an attemp to visualize the interconnection between the different formulations can be seen in Figure 1.2.2. The formulations within this diagram are classified by (in order)

- the set of unknowns [green]

- the differential operator acting on the Finite Element shape functions [blue]

- the constraints imposed on the formulation (explicitly, through the space or through lagrange multipliers)[red]

- spaces (continuous/discontinous/normal-continuous) for u (and σ) [black]

	primal	dual	primal DG	dual DG
Standard	CG FEM u ∇u $[\![\mathbf{u}]\!]=0$ u is cont.		DG FEM u ∇u $[\![\hat{\sigma}]\!]_n=[\![\hat{u}]\!]=0$ u discont.	
mixed		M FEM $(u,\underline{\sigma})$ div($\underline{\sigma}$) $[\![\sigma]\!]_n=0$ u disc., σ_n cont.	DG FEM $(u,\underline{\sigma})$ ∇u $[\![\hat{\sigma}]\!]_n=[\![\hat{u}]\!]=0$ u,σ discont.	DG FEM $(u,\underline{\sigma})$ div($\underline{\sigma}$) $[\![\hat{\sigma}]\!]_n=[\![\hat{u}]\!]=0$ u,σ discont.
hybrid		HM FEM $(u,\underline{\sigma},u_F)$ div($\underline{\sigma}$) $[\![\sigma]\!]_n=0$ u,σ discont.	HDG FEM (u,u_F) ∇u $[\![\hat{\sigma}]\!]_n=[\![u_F]\!]=0$ u discont.	HDG FEM $(u,\underline{\sigma},u_F)$ div($\underline{\sigma}$) $[\![\hat{\sigma}]\!]_n=[\![u_F]\!]=0$ u,σ discont.
		abbreviations: C: continuous D: discontinuous M: mixed H: hybrid G: Galerkin	HDG FEM $(u,\sigma_{n,F})$ ∇u $[\![\sigma_{n,F}]\!]=[\![\hat{u}]\!]=0$ u discont.	
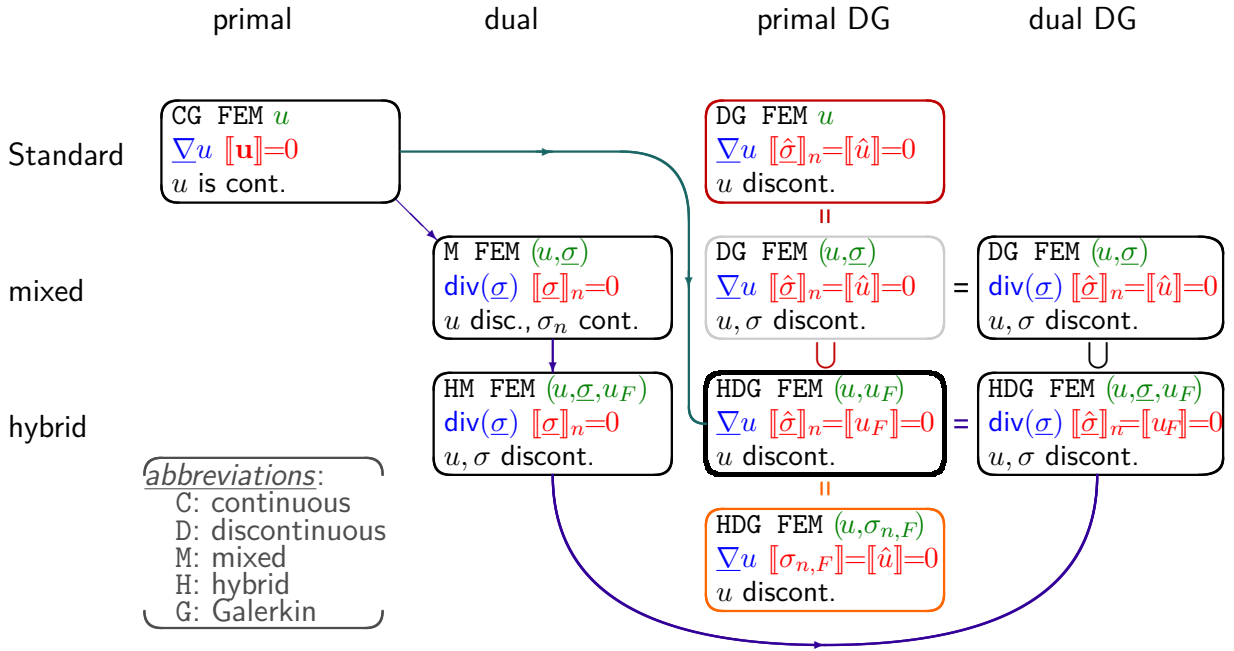
Figure 1.2.2: Interconnection of several Finite Element methods

This diagram shows the relations of the proposed HDG method (bold) to other methods. Whereas we went the direct way from CG FEM to our formulation [turquoise path] when introducing the formulation, we also pointed out the relation to other methods in the last section. Therefore we presented the HDG method as

1.2.2.1 a hybridized mixed method with a modification of the conservation constraint ($[\![\underline{\sigma}]\!]_n = 0 \rightarrow [\![\hat{\sigma}]\!] = 0$). We went the way from CG FEM to HDG FEM in dual DG form over mixed FEM (M FEM) and hybridized mixed FEM (HM FEM) [purple path].

1.2.2.2 a Standard DG Formulation without additional unknowns u_F. It was shown that HDG FEM is a subset of DG FEM [dark red] by eliminating the lagrange multiplier u_F and presenting the equivalent Standard DG formulation.

1.2.2.3 a stabilized DG formulation. Here we showed another equivalent representation of the proposed method which uses additional unknowns $\sigma_{n,F}$ for the flux - instead of the variable u_F - on the facets [orange].

1.2.3   A priori error analysis

For a precise analysis, we have to define appropriate spaces to work with. In contrast to CG our discrete variational formulation includes normal derivates which are not defined for all function in $H^1(\Omega)$. So we have to use additional regularity. We assume our problem has a solution $u \in H^1(\Omega) \cap H^2(\mathcal{T}_h) =: W$. Remember that we already defined $V = W \times L^2(\mathcal{F}_h)$, s.t. our bilinearform $\mathcal{B}_h : V \times V \to \mathbb{R}$ is well-defined.

For our a priori error estimates we will first show consistency and adjoint consistency of our method, which will later on give us Galerkin orthogonality on the one hand and will allow us for using *Aubin-Nitsche* arguments to get error estimates in the L^2-norm on the other hand. Then we'll show coercivity and boundedness of the bilinearform with the respect to an HDG norm. These results will then be used for the error estimates, but they already imply uniquely solvability of the discrete problem itself. The last ingredient will be the approximation property of the chosen Finite Element space in the HDG norm. Then we can put all those results together to show optimal a priori error estimates for the HDG method.

Before we start with the analysis lets summarize the introduced spaces:

$$W = H^1(\Omega) \cap H^2(\mathcal{T}_h) \tag{1.2.35a}$$
$$V = \left\{ (u, u_F) : u \in H^2(\mathcal{T}_h) \cap H^1(\Omega), u_F \in L^2(\mathcal{F}_h) \right\} = W \times L^2(\mathcal{F}_h) \tag{1.2.35b}$$
$$V_h^k = \left\{ (u, u_F) : u \in \mathcal{P}^k(T) \ \forall \ T \in \mathcal{T}_h, u_F \in \mathcal{P}^k(E) \ \forall \ E \in \mathcal{F}_h \right\} = Q_h^k \times F_h^k \subset V \tag{1.2.35c}$$

Additional subscripts 0 and D indicate homogeneous and nonhomogeneous Dirichlet conditions which are incorporated into the space in an L^2-projective sense. Attention: $V_{h,0} \subset V_0$ but $V_{h,D} \nsubseteq V_D$, i.e. L^2-projected Dirichlet data may not coincide with exact Dirichlet data.

The superscripted k which describes the used (uniform) polynomial order of the space, will often be dropped. During our analysis we always assume that the solution u of (1.2.1) lies in W and ε is element-wise continuous. For simplicity we furthermore assume that we have only Dirichlet boundary conditions on the whole boundary $\partial\Omega$. Modifications for Neumann and Robin boundary conditions (even without Dirichlet boundary conditions) are trivial.

In order to make $u \in W$ compatible with functions in V we introduce the identity operator

$$\mathbf{id}_h : W \to V, \qquad \mathbf{id}_h u = (u, \ \mathrm{tr}|_{\mathcal{F}_h}(u))$$

The subindex h indicates that the identity operator depends on the mesh as $\mathcal{F}_h$ does. We will use this identity operator at several points implicitly by writing $\mathbf{u}$ for $\mathbf{id}_h u$. Only when additional emphasis on the different spaces seems necessary we will use this operator explicitly.

At several places we will work with a "characteristic" size of an element, which will be denoted by h. As for well shaped elements different choices of h are possible and those differ only by a constant of order one, we won't determine the choice of this element size h further. E.g., think of it as the diameter of the element.

Remark 1.2.5 (Variable order):
Throughout the analysis in this work we assume uniform polynomial order, but nevertheless, the derivation of modified results is straight forward.

1.2.3.1   Consistency

Our aim here is to show that our Bilinearform is consistent, such that Galerkin-Orthogonality holds. So we remember that $\mathrm{tr}|_E(u)$ is single-valued for u in W, s.t. $[\![\mathbf{u}]\!] = 0$. Let $\mathbf{u}$ be the solution of

(1.2.1), then for any $\mathbf{v} = (v, v_F) \in V_0$ we get with the definition of the bilinearform (1.2.8):

$$
\begin{aligned}
\mathcal{B}_h(\mathbf{u}, \mathbf{v}) &= \sum_{T \in \mathcal{T}_h} \left\{ \int_T \varepsilon \, \underline{\nabla} u \, \underline{\nabla} v \, d\underline{x} - \int_{\partial T} \varepsilon \, \frac{\partial u}{\partial \underline{n}} \, (v - v_F) \, d\underline{s} - \int_{\partial T} \underbrace{[\![u]\!]}_{=0} \varepsilon \left(\frac{\partial v}{\partial \underline{n}} - \tau_h \, [\![\mathbf{v}]\!] \right) d\underline{s} \right\} \\
&= \sum_{T \in \mathcal{T}_h} \left\{ \int_T -\mathrm{div}(\varepsilon \, \underline{\nabla} u) \, v \, d\underline{x} + \int_{\partial T} \varepsilon \, \frac{\partial u}{\partial \underline{n}} \, v_F \, d\underline{s} \right\} \\
&= \int_\Omega f \, v \, d\underline{x} + \sum_{E \in \mathcal{F}_h^{int}} \int_E \underbrace{[\![\varepsilon \, \underline{\nabla} u]\!]_n}_{=0} v_F \, d\underline{s} + \sum_{E \in \mathcal{F}_h^{ext}} \int_E \varepsilon \, \frac{\partial u}{\partial \underline{n}} \, \underbrace{v_F}_{=0} \, d\underline{s}
\end{aligned}
\tag{1.2.36}
$$

In the last step we used that the integrals from both sides of an interior facet can be summed up within one integral. Furthermore we made use of the relation $[\![\varepsilon \, \underline{\nabla} u]\!] = 0$, which holds true since u is the solution of the PDE (1.2.1). As the problem and the bilinearform are symmetric, we also directly get consistency of the adjoint problem, which will allow us for using Aubin-Nitsche arguments for L^2-norm error estimates later:

$$
\mathcal{B}_h(\mathbf{u}, \mathbf{v}) = \mathcal{B}_h(\mathbf{v}, \mathbf{u}) = \int_\Omega f \, v \, d\underline{x} \qquad \forall \, \mathbf{v} \in V_0
\tag{1.2.37}
$$

Proposition 1.2.1 (Galerkin Orthogonality). *Let $\mathbf{u}_h \in V_{h,D}$ be the solution of* (1.2.9) *and $\mathbf{u}$ be the solution of* (1.2.1). *Then there holds*

$$
\mathcal{B}_h(\mathbf{u}_h - \mathbf{u}, \mathbf{v}_h) = 0 \qquad \forall \, \mathbf{v}_h \in V_{h,0}
\tag{1.2.38}
$$

Proof. Due to (1.2.36) there holds

$$
\mathcal{B}_h(\mathbf{u}_h - \mathbf{u}, \mathbf{v}_h) = \mathcal{B}_h(\mathbf{u}_h, \mathbf{v}_h) - \mathcal{B}_h(\mathbf{u}, \mathbf{v}_h) = \int_\Omega f \, v_h \, d\underline{x} - \int_\Omega f \, v_h \, d\underline{x} = 0 \qquad \forall \, \mathbf{v}_h \in V_{h,0} \tag{1.2.39}
$$

$\square$

Note that the restriction to Dirichlet boundary condition is not at all necessary here, but was only chosen for simplicity.

1.2.3.2 Stability and boundedness

We assume that ε and τ_h are constant on each element and facet respectively. Due to the non-conformity of our method, i.e. due to the fact that solutions of the discrete problem are not in $H^1(\Omega)$, we can not work with the $H^1(\Omega)$ (semi-) norm to show stability and boundedness of the introduced bilinearform. Instead, we have to define new norms. An appropriate discrete norm for the broken Sobolev space $H^1(\mathcal{T}_h)$, which is the natural one to show coercivity in, is

$$
\|\mathbf{u}\|_{\varepsilon,*}^2 := \sum_{T \in \mathcal{T}_h} \left\{ \varepsilon \|\underline{\nabla} u\|_T^2 + \frac{\varepsilon}{h} [\![\mathbf{u}]\!]_{\partial T}^2 \right\}
\tag{1.2.40}
$$

where $[\![\cdot]\!]_{\partial T} := \|[\![\cdot]\!]\|_{\partial T}$.
This is actually a semi-norm similar to the H^1 semi-norm for conforming Finite Elements. But as the broken Sobolev space $H^1(\mathcal{T}_h)$ does not contain any continuity restrictions on element interfaces the jump terms appear in the norm.

The kernel of the norm operator is the space of constant functions. If we now pose Dirichlet conditions on a (nonzero) part of the boundary or constrain an integral average of u $\int_G u\,d\underline{x} = C \in \mathbb{R}$ for G some nonzero part of the domain the kernel is only the trivial one and the proposed operator is actually a norm[8].

In this norm we can show coercivity of $\mathcal{B}_h : V_h \times V_h \to \mathbb{R}$. But as we also have to show boundedness of $\mathcal{B}_h$ for $v \in V_h$ and $u \in V$ we have to consider another norm to bound normal derivative terms by it. Therefore we introduce the second *HDG norm*

$$\|\mathbf{u}\|_\varepsilon^2 := \sum_{T \in \mathcal{T}_h} \left\{ \varepsilon \|\underline{\nabla} u\|_T^2 + \frac{\varepsilon}{h} [\![\mathbf{u}]\!]_{\partial T}^2 + \varepsilon h \|\frac{\partial u}{\partial \underline{n}}\|_{\partial T}^2 \right\} \tag{1.2.41}$$

which is the natural one to boundedness of the bilinearform in.

For $\mathbf{u} \in V_h$ the norms $\|\mathbf{u}\|_\varepsilon$ and $\|\mathbf{u}\|_{\varepsilon,*}$ are equivalent. As $\mathbf{u} \in V_h$ is piecewise polynomial we can bound the difference of both norms by $\|\mathbf{u}\|_{\varepsilon,*}^2$ with the inverse inequality (see (A.2.3b)):

$$\|\mathbf{u}\|_\varepsilon^2 - \|\mathbf{u}\|_{\varepsilon,*}^2 = \sum_{T \in \mathcal{T}_h} \varepsilon h \|\frac{\partial u}{\partial \underline{n}}\|_{\partial T}^2 \le \sum_{T \in \mathcal{T}_h} \varepsilon c_T^2 \|\underline{\nabla} u\|_T^2 \le c_{sr}^2 \|\mathbf{u}\|_{\varepsilon,*}^2 \tag{1.2.42}$$

with $c_{sr} = \max\limits_{T \in \mathcal{T}_h} c_T \in \mathbb{R}$ only depending on the shape and not on the size of the elements. Obviously then there holds:

$$\frac{1}{\sqrt{1 + c_{sr}^2}} \|\mathbf{u}\|_\varepsilon \le \|\mathbf{u}\|_{\varepsilon,*} \le \|\mathbf{u}\|_\varepsilon \qquad \forall\, \mathbf{u} \in V_h \tag{1.2.43}$$

We will denote each element contribution by $\|\!|\cdot\|\!|^T$ and $\|\!|\cdot\|\!|^T$.

Coercivity

With respect to those norms we can turn over to coercivity and stability. Let's start with coercivity:

Proposition 1.2.2 (Coercivity). *For a shape regular mesh and $\tau_h h$ (with h the local mesh size) sufficiently large $\mathcal{B}_h(\cdot, \cdot)$ is coercive on V_h with respect to the norm $\|\!|\cdot\|\!|_\varepsilon$, that is*

$$\mathcal{B}_h(\mathbf{u}, \mathbf{u}) \ge c \|\mathbf{u}\|_{\varepsilon,*}^2 \ge \alpha_{\mathcal{B}_h} \|\mathbf{u}\|_\varepsilon^2 \qquad \forall\, \mathbf{u} \in V_h \tag{1.2.44}$$

with c, $\alpha_{\mathcal{B}_h} \in \mathbb{R}$ independent of the mesh size.

Proof. We will make use of several elementary inequalities which are listed in section A.2 of the appendix. These are Cauchy Schwarz inequalities, Young's inequality and inverse inequalities. With

---

[8]If neither is true and $\Gamma_D = \varnothing$ and $\Gamma_R \ne \varnothing$ add $\frac{\varepsilon}{h}\|u\|_{\Gamma_R}^2$ to $\|\mathbf{u}\|_{\varepsilon,*}^2$ and you have a norm. All following results can also be adapted very easily to this modified norm

them we can directly prove the Proposition. For all $\mathbf{u} \in V_h$ there holds

$$
\begin{aligned}
\mathcal{B}_h(\mathbf{u}, \mathbf{u}) \quad &= \quad \sum_{T \in \mathcal{T}_h} \varepsilon \left\{ \|\underline{\nabla} u\|_T^2 \quad - \quad 2 \left(\frac{\partial u}{\partial \underline{n}}, [\![\mathbf{u}]\!] \right)_{\partial T} \quad + \quad \tau_h [\![\mathbf{u}]\!]_{\partial T}^2 \right\} \\
&\overset{(A.2.1a)}{\geq} \quad \sum_{T \in \mathcal{T}_h} \varepsilon \left\{ \|\underline{\nabla} u\|_T^2 \quad - \quad 2 \left\| \frac{\partial u}{\partial \underline{n}} \right\|_{\partial T} [\![\mathbf{u}]\!]_{\partial T} \quad + \quad \tau_h [\![\mathbf{u}]\!]_{\partial T}^2 \right\} \\
&\overset{(A.2.3b)}{\geq} \quad \sum_{T \in \mathcal{T}_h} \varepsilon \left\{ \|\underline{\nabla} u\|_T^2 \quad - \quad \|\underline{\nabla} u\|_T \left(2 \frac{c_T}{\sqrt{h}} [\![\mathbf{u}]\!]_{\partial T} \right) \quad + \quad \tau_h [\![\mathbf{u}]\!]_{\partial T}^2 \right\} \\
&\overset{(A.2.2)}{\geq} \quad \sum_{T \in \mathcal{T}_h} \varepsilon \left\{ \|\underline{\nabla} u\|_T^2 \quad - \quad \frac{1}{2} \|\underline{\nabla} u\|_T^2 - \frac{1}{2} \left(\frac{2c_T}{\sqrt{h}} [\![\mathbf{u}]\!]_{\partial T} \right)^2 \quad + \quad \tau_h [\![\mathbf{u}]\!]_{\partial T}^2 \right\} \\
&= \quad \sum_{T \in \mathcal{T}_h} \varepsilon \left\{ \frac{1}{2} \|\underline{\nabla} u\|_T^2 \quad + \quad \left(\tau_h - \frac{2c_T^2}{h} \right) [\![\mathbf{u}]\!]_{\partial T}^2 \right\} \\
&\geq \quad \min\left(\frac{1}{2}, \tau_h h - 2c_{sr}^2 \right) \|\mathbf{u}\|_{\varepsilon,*}^2 \\
&\overset{(1.2.43)}{\geq} \quad \frac{1}{\sqrt{1+c_{sr}^2}} \min\left(\frac{1}{2}, \tau_h h - 2c_{sr}^2 \right) \|\mathbf{u}\|_\varepsilon^2 \\
&\geq \quad \frac{1}{2\sqrt{1+c_{sr}^2}} \|\mathbf{u}\|_\varepsilon^2
\end{aligned}
\tag{1.2.45}
$$

where in the last step we used $\tau_h \geq \frac{2c_{sr}^2 + \frac{1}{2}}{h}$. So with $c = \frac{1}{2}$ and $\alpha_{\mathcal{B}_h} = \frac{1}{2\sqrt{1+c_{sr}^2}}$ the claim holds true. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Coercivity implies uniqueness of the discrete solution as for two discrete solutions $\mathbf{u}_1, \mathbf{u}_2 \in V_{h,D}$ there holds

$$
0 = \mathcal{B}_h(\mathbf{u}_1 - \mathbf{u}_2, \mathbf{u}_1 - \mathbf{u}_2) \geq C \|\mathbf{u}_1 - \mathbf{u}_2\|_\varepsilon
$$

and as $\|\cdot\|_\varepsilon$ is a norm for $V_{h,0}$ and $\mathbf{u}_1 - \mathbf{u}_2$ is zero on dirichlet boundaries $\mathbf{u}_1$ and $\mathbf{u}_2$ have to coincide.

Remark 1.2.6 (Stabilization parameter):
As we see in the appendix the constant of the inverse inequality c_{sr} or locally c_T depends linearly on the polynomial degree k. So with the condition $\tau_h > \frac{2c_{sr}^2 + \frac{1}{2}}{h}$, it is sufficient to fulfill $\tau_h > \frac{\alpha_0}{h} k^2$ for an α_0 only depending on the shape. It is of course sufficient if those relations are fulfilled locally $(c_{sr} \to c_T)$.

Boundedness
To show boundedness of the bilinearform in the $\|\cdot\|_\varepsilon$-norm is even easier, as we will see now:

Proposition 1.2.3 (Boundedness). *For all $\mathbf{u}, \mathbf{v} \in V$ there holds:*

$$
|\mathcal{B}_h(\mathbf{u}, \mathbf{v})| \leq \beta_{\mathcal{B}_h} \|\mathbf{u}\|_\varepsilon \|\mathbf{v}\|_\varepsilon
$$

with $\beta_{\mathcal{B}_h} = \sup_{\underline{x} \in \mathcal{F}_h}(1 + \tau_h h)$

Proof. We can prove the claim with a combination of different versions of the Cauchy-Schwarz in-

equality:

$$
\begin{aligned}
&|\mathcal{B}_h(\mathbf{u}, \mathbf{v})| \\
&= \left| \sum_{T \in \mathcal{T}_h} \varepsilon \left\{ (\underline{\nabla} u, \underline{\nabla} v)_T \quad - \quad (\frac{\partial u}{\partial \underline{n}}, [\![\mathbf{v}]\!])_{\partial T} \quad - \quad (\frac{\partial v}{\partial \underline{n}}, [\![\mathbf{u}]\!])_{\partial T} \quad + \quad \tau_h([\![\mathbf{u}]\!], [\![\mathbf{v}]\!])_{\partial T} \right\} \right| \\
&\overset{(A.2.1a)}{\leq} \sum_{T \in \mathcal{T}_h} \varepsilon \left\{ \|\underline{\nabla} u\|_T \|\underline{\nabla} v\|_T + \left\|\frac{\partial u}{\partial \underline{n}}\right\|_{\partial T} [\![\mathbf{v}]\!]_{\partial T} + \left\|\frac{\partial v}{\partial \underline{n}}\right\|_{\partial T} [\![\mathbf{u}]\!]_{\partial T} + \tau_h [\![\mathbf{u}]\!]_{\partial T} [\![\mathbf{v}]\!]_{\partial T} \right\} \\
&\overset{(A.2.1b)}{\leq} \sum_{T \in \mathcal{T}_h} \varepsilon \left\{ \left(\|\underline{\nabla} u\|_T^2 \quad + \quad h \left\|\frac{\partial u}{\partial \underline{n}}\right\|_{\partial T}^2 \quad + \quad \frac{1}{h}[\![\mathbf{u}]\!]_{\partial T}^2 \quad + \quad \tau_h [\![\mathbf{u}]\!]_{\partial T}^2 \right)^{\frac{1}{2}} \right. \\
&\qquad\qquad \left. \cdot \left(\|\underline{\nabla} v\|_T^2 \quad + \quad h \left\|\frac{\partial v}{\partial \underline{n}}\right\|_{\partial T}^2 \quad + \quad \frac{1}{h}[\![\mathbf{v}]\!]_{\partial T}^2 \quad + \quad \tau_h [\![\mathbf{v}]\!]_{\partial T}^2 \right)^{\frac{1}{2}} \right\} \\
&\overset{(A.2.1b)}{\leq} \left(\sum_{T \in \mathcal{T}_h} \varepsilon \left\{ \|\underline{\nabla} u\|_T^2 \quad + \quad h \left\|\frac{\partial u}{\partial \underline{n}}\right\|_{\partial T}^2 \quad + \quad \frac{1}{h}[\![\mathbf{u}]\!]_{\partial T}^2 \quad + \quad \tau_h [\![\mathbf{u}]\!]_{\partial T}^2 \right\} \right)^{\frac{1}{2}} \\
&\qquad\qquad \cdot \left(\sum_{T \in \mathcal{T}_h} \varepsilon \left\{ \|\underline{\nabla} v\|_T^2 \quad + \quad h \left\|\frac{\partial v}{\partial \underline{n}}\right\|_{\partial T}^2 \quad + \quad \frac{1}{h}[\![\mathbf{v}]\!]_{\partial T}^2 \quad + \quad \tau_h [\![\mathbf{v}]\!]_{\partial T}^2 \right\} \right)^{\frac{1}{2}} \\
&= \left(\|\mathbf{u}\|_\varepsilon^2 + \sum_{T \in \mathcal{T}_h} \varepsilon \tau_h [\![\mathbf{u}]\!]_{\partial T}^2 \right)^{\frac{1}{2}} \cdot \left(\|\mathbf{v}\|_\varepsilon^2 + \sum_{T \in \mathcal{T}_h} \varepsilon \tau_h [\![\mathbf{v}]\!]_{\partial T}^2 \right)^{\frac{1}{2}} \\
&\overset{\max}{\leq} \sup_{\underline{x} \in \mathcal{F}_h} (1 + \tau_h h) \ \|\mathbf{u}\|_\varepsilon \ \|\mathbf{v}\|_\varepsilon
\end{aligned}
$$

(1.2.46)

$\square$

Remark 1.2.7 (non-constant coefficients):
Notice that the assumption of piecewise constant coefficients is not at all necessary for proving stability or boundedness. If you interprete $\varepsilon\|\underline{\nabla} u\|_T^2$ and $\frac{\varepsilon}{h}[\![\mathbf{u}]\!]_{\partial T}^2$ as short notations for $(\varepsilon \underline{\nabla} u, \underline{\nabla} u)_T$ and $(\frac{\varepsilon}{h}[\![\mathbf{u}]\!], [\![\mathbf{u}]\!])_{\partial T}$ you directly get both inequalities in the more general case. The only difference makes the inverse inequality (used in the step from line two to line three in equation (1.2.45)). Here, you would have to adjust the constant c_T to the properties you request on the coefficient ε.

1.2.3.3   Approximation

Let's restate one version of the famous *Bramble Hilbert Lemma*

Lemma 1.2.4 (Bramble Hilbert Lemma). *Let U be some Hilbert space, and $L : H^m \rightarrow U$, $m = k+1$ be a continuous linear operator such that $Lq = 0$ for all polynomials $q \in \mathcal{P}^k$. Then there holds*

$$
\|Lv\|_U \leq C|v|_{H^m} \qquad \forall \, v \in H^m
$$

with C independent of u and the size of the domain.

Now we introduce the element local L^2-projector $\mathbf{\Pi}_{h,k} : H^m(T) \to \mathcal{P}^k(T) \times F_h^k(T)$ which is defined by the following relations

$$\mathbf{\Pi}_{h,k}(u) := \left(\Pi_{h,k}^{el}(u), \ \Pi_{h,k}^{fac}(u) \right) \quad \text{with}$$

$$
\begin{aligned}
\left(\Pi_{h,k}^{el}(u), \ v \right)_T &= (u,v)_T & \forall \ v \ \in \mathcal{P}^k(T) \\
\left(\Pi_{h,k}^{fac}(u), v_F \right)_{\partial T} &= \left(\mathrm{tr}|_{\mathcal{F}_h}(u), v_F \right)_{\partial T} & \forall \ v_F \in F_h^k(T)\mathcal{P}^k(E)
\end{aligned}
\tag{1.2.47}
$$

With those preparations we can now state and prove the following Lemma.

Lemma 1.2.5. *For $u \in H^m(T)$, $m = k+1$, $k \geq 1$ and the L^2-projector $\mathbf{\Pi}_{h,k} : H^m(T) \to \mathcal{P}^k(T) \times F_h^k(T)$ of (1.2.47), there holds*

$$\|(\mathbf{id}_h - \mathbf{\Pi}_{h,k})u\|_\varepsilon^T \leq \sqrt{\|\varepsilon\|}_\infty \, C \, |u|_{H^m(T)} \qquad \forall \ u \in H^m(T) \tag{1.2.48}$$

with C independent of u and element size.

Proof. We use the *Bramble Hilbert Lemma* with $U = H^2(T) \times L^2(\partial T)$ and the scalar product

$$(\mathbf{u}, \mathbf{v})_U := \int_T \underline{\nabla} u \, \underline{\nabla} v \, d\underline{x} + \int_{\partial T} \frac{1}{h} [\![\mathbf{u}]\!][\![\mathbf{v}]\!] \, d\underline{s} + \int_{\partial T} h \frac{\partial u}{\partial \underline{n}} \frac{\partial v}{\partial \underline{n}} \, d\underline{s} + \underbrace{\frac{1}{h^2} \left(\int_{\partial T} u \, d\underline{s} \right) \left(\int_{\partial T} v \, d\underline{s} \right)}_{=:R(\mathbf{u},\mathbf{v})}$$

such that $(\mathbf{u}, \mathbf{u})_U = \|\mathbf{u}\|_U^2 = (\|\frac{1}{\sqrt{\varepsilon}}\mathbf{u}\|_\varepsilon^T)^2 + R(\mathbf{u}, \mathbf{v})$. The operator L which is chosen as

$$L := \mathbf{id}_h - \mathbf{\Pi}_{h,k}$$

is clearly linear and for all polynomials $q \in \mathcal{P}^k$ there holds $Lq = 0$. We still have to show that L is a continuous operator. Note that for $u \in H^2(T)$ the trace of u and the trace of the normal derivative are continuous linear operators such that $\|\mathbf{u}\|_U = \|\frac{1}{\sqrt{\varepsilon}}\mathbf{u}\|_\varepsilon^T + \sqrt{R(\mathbf{u}, \mathbf{u})} \leq C\|u\|_{H^2}$ holds. Furthermore the introduced L^2-projector $(w, w_F) = \mathbf{\Pi}_{h,k}u$ of a function u onto the polynomial space $\mathcal{P}^k(T) \times F_h^k(T)$ can be bounded in the $\|\cdot\|_U$-norm by the $H^1(T)$-norm of w and the $L^2(\partial T)$-norm of w_F.
As a consequence of both, the U-norm of Lu can be bounded in the $H^2(T)$-norm of u as well

$$\|Lu\|_U \leq \|\mathbf{id}_h u\|_U + \|\mathbf{\Pi}_{h,k}u\|_U \leq C\|u\|_{H^2(T)} \tag{1.2.49}$$

So all requirements of the *Bramble Hilbert Lemma* are fulfilled and we obtain

$$\|(\mathbf{id}_h - \mathbf{\Pi}_{h,k})u\|_\varepsilon^T \leq \|\sqrt{\varepsilon}Lu\|_U \leq \sqrt{\|\varepsilon\|}_\infty \, C \, |u|_{H^m(T)} \qquad \forall \ u \in H^m(T) \tag{1.2.50}$$

$\square$

Now we will look at the global interpolation error to bound the *approximation error* $\inf_{\mathbf{v}_h \in V_h} \|\mathbf{u} - \mathbf{v}_h\|_\varepsilon$. Therefore we will transform all element contributions on a reference element. There we will use Lemma 1.2.5 and by transforming back we will get the correct h scaling of the interpolation error.

Proposition 1.2.6. *On a shape regular mesh $\mathcal{T}_h$ which consists of affine transformed elements and a function $u \in H^1(\Omega) \cap H^m(\mathcal{T}_h)$, $m = k + 1 \geq 2$ there holds the following approximation result for V_h^k as introduced in* (1.2.35c)

$$\inf_{\mathbf{v}_h \in V_h^k} \|\mathbf{u} - \mathbf{v}_h\|_\varepsilon^2 \leq C \sum_{T \in \mathcal{T}_h} \|\varepsilon\|_\infty \, h_T^{2k} \, |u|_{H^m(T)}^2 \tag{1.2.51}$$

with C independent of the mesh size. For quasi-uniform meshes the direct conclusion of it is

$$\inf_{\mathbf{v}_h \in V_h^k} \|\mathbf{u} - \mathbf{v}_h\|_\varepsilon \leq C \, \sqrt{\|\varepsilon\|_\infty} \, h^k \, |u|_{H^m(\mathcal{T}_h)} \tag{1.2.52}$$

Proof. Let's first note that as $u \in H^1(\Omega)$ the trace of u onto facets is single-valued, such that the L^2-projected approximation is also single-valued. So we are able to define the global projection-operator $\mathbf{\Pi}_{h,k}^\Omega : H^m(\mathcal{T}_h) \to V_h$ as an element-wise projection with $\mathbf{\Pi}_{h,k}^T$ the L^2-projector of (1.2.47) with respect to a single element T. So we start with

$$\inf_{\mathbf{v}_h \in V_h^k} \|\mathbf{u} - \mathbf{v}_h\|_\varepsilon^2 \leq \|\mathbf{u} - \mathbf{\Pi}_{h,k}^\Omega u\|_\varepsilon^2 = \sum_{T \in \mathcal{T}_h} \left(\|(\mathbf{id}_h^T - \mathbf{\Pi}_{h,k}^T)u\|_\varepsilon^T \right)^2 \tag{1.2.53a}$$

where $\mathbf{id}_h^T u$ is the restriction of $\mathbf{id}_h u$ to the element T. Now we can transform each element T to an according reference element $\hat{T}$ of size $\mathcal{O}(1)$, where the transformation from reference element $\hat{T}$ to physical element T is defined by $F_T : \hat{T} \to T$.
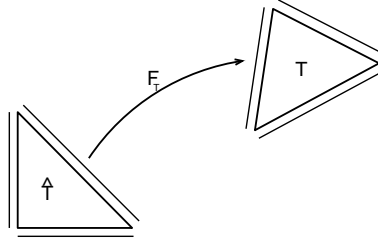
Figure 1.2.3: Affine transformation F_T from reference domain $\hat{T}$ to the physical domain T

On each element we then get

$$\left(\|(\mathbf{id}_h^T - \mathbf{\Pi}_{h,k}^T)u\|_\varepsilon^T \right)^2 \preceq h_T^d h_T^{-2} \left(\|(\mathbf{id}_h^T - \mathbf{\Pi}_{h,k}^T)u \circ F_T\|_\varepsilon^{\hat{T}} \right)^2 \tag{1.2.53b}$$

where we used that the scaling of $\|\mathbf{v}\|_\varepsilon$ is h_T^{d-2}, which becomes clear when we look at the three contributions of the norm separately:

part of $\|\cdot\|_\varepsilon^2$	scaling factor	integral trafo.	diff. op. scaling	overall scaling
$\int_T \varepsilon(\nabla v)^2 \, d\underline{x}$	1	$\simeq h_T^d$	$\simeq h_T^{-2}$	$\simeq h_T^{d-2}$
$\dfrac{1}{h}\int_{\partial T} \varepsilon[\![\mathbf{v}]\!]^2 \, d\underline{s}$	h_T^{-1}	$\simeq h_T^{d-1}$	1	$\simeq h_T^{d-2}$
$h \int_{\partial T} \varepsilon \dfrac{\partial v}{\partial \underline{n}}^2 \, d\underline{s}$	h_T	$\simeq h_T^{d-1}$	$\simeq h_T^{-2}$	$\simeq h_T^{d-2}$

Here, we can use another property of the L^2-projection: it commutes with the transformation, i.e. the result of the L^2-projection of a function transformed from the reference to the physical domain and the transformation from reference to physical domain of the L^2-projection of a function are the same. So there also holds:

$$\left(\||(\mathbf{id}_h^T - \mathbf{\Pi}_{h,k}^T)u\||_\varepsilon^T\right)^2 \quad \preceq \quad h_T^d \, h_T^{-2} \left(\||(\mathbf{id}_h^{\hat{T}} - \mathbf{\Pi}_{h,k}^{\hat{T}})(u \circ F_T)\||_\varepsilon^{\hat{T}}\right)^2 \tag{1.2.53c}$$

We finally reached the point where Lemma 1.2.5 can be applied on each element to get

$$\left(\||(\mathbf{id}_h^T - \mathbf{\Pi}_{h,k}^T)u\||_\varepsilon^T\right)^2 \quad \preceq \quad \|\varepsilon\|_\infty h_T^d \, h_T^{-2} \, |u \circ F_T|_{H^m(\hat{T})}^2 \tag{1.2.53d}$$

Now we transform back to physical domain and achieve

$$\left(\||(\mathbf{id}_h^T - \mathbf{\Pi}_{h,k}^T)u\||_\varepsilon^T\right)^2 \quad \preceq \quad \|\varepsilon\|_\infty h_T^d \, h_T^{-2} \, h_T^{-d} \, h_T^{2m} \, |u|_{H^m(T)}^2 = \|\varepsilon\|_\infty h_T^{2k} \, |u|_{H^m(T)}^2 \tag{1.2.53e}$$

This proves the first part of the proposition. If we consider quasi-uniform meshes we get $h \simeq h_T \, \forall \, T \in \mathcal{T}_h$ and can move the factor h_T^{2k} in front of the sum, what proves the second part of the proposition.   □

Remark 1.2.8 (Boundary conditions):
We did not use special treatment for boundary conditions as we assume Dirichlet boundary conditions that are incorporated into the space are either exact or approximations in an L^2-orthogonal sense. In both cases V_h can be exchanged by $V_{h,D}$ and the results stay uneffected. If neither is true, the changes in the result are small as long as the error of approximation is of the same magnitude as the remainder.

1.2.3.4   Putting it all together

HDG norm estimates
First we are interested in error estimates in the HDG norm of the form

$$\|\mathbf{u} - \mathbf{u}_h\|_\varepsilon \leq Ch^q|u|_{H^m(\mathcal{T}_h)} \tag{1.2.54}$$

with q depending on the chosen polynomial approximation and the regularity of the weak problem. Before we can show this kind of estimate we need the following Lemma:

Lemma 1.2.7 (Modified Cea's Lemma). *Let $\mathbf{u}_h \in V_h^k$ be the solution of* (1.2.9) *and $u \in H^1(\Omega) \cap H^m(\mathcal{T}_h) \subset W$, $m \geq 2$ the solution of* (1.2.1). *Then the error is only bounded a constant away from the* approximation error:

$$\|\mathbf{u} - \mathbf{u}_h\|_\varepsilon \leq C \inf_{\mathbf{v}_h \in V_{h,D}} \|\mathbf{u} - \mathbf{v}_h\|_\varepsilon \tag{1.2.55}$$

for $C \in \mathbb{R}$ only depending on τ_h and the shape regularity.

Proof. We start with a triangle inequality to separate an expression that is later on going to be the *discrete error* and one that is going to be the *approximation error*

$$\|\mathbf{u} - \mathbf{u}_h\|_\varepsilon^2 \leq \|\mathbf{u} - \mathbf{v}_h\|_\varepsilon^2 + \|\mathbf{v}_h - \mathbf{u}_h\|_\varepsilon^2 \qquad \forall \, \mathbf{v}_h \in V_{h,D} \tag{1.2.56}$$

From section 1.2.3.3 we have the estimates for the *approximation error*. Now we need estimates for the other part $\|\mathbf{v}_h - \mathbf{u}_h\|_\varepsilon$, which we easily get from coercivity, Galerkin orthogonality and boundedness of the bilinearform:

$$
\begin{aligned}
\|\mathbf{v}_h - \mathbf{u}_h\|_\varepsilon^2 &\overset{(1.2.45)}{\leq} \frac{1}{\alpha_{\mathcal{B}_h}} \mathcal{B}_h(\mathbf{v}_h - \mathbf{u}_h, \underbrace{\mathbf{v}_h - \mathbf{u}_h}_{\mathbf{w}_h \in V_{h,0}}) \\
&\overset{(1.2.38)}{=} \frac{1}{\alpha_{\mathcal{B}_h}} \left(\mathcal{B}_h(\mathbf{v}_h - \mathbf{u}, \mathbf{v}_h - \mathbf{u}_h) + \underbrace{\mathcal{B}_h(\mathbf{u} - \mathbf{u}_h, \mathbf{w}_h)}_{=0} \right) \\
&\overset{(1.2.46)}{\leq} \frac{\beta_{\mathcal{B}_h}}{\alpha_{\mathcal{B}_h}} \|\mathbf{v}_h - \mathbf{u}\|_\varepsilon \cdot \|\mathbf{v}_h - \mathbf{u}_h\|_\varepsilon
\end{aligned}
\tag{1.2.57}
$$

Thus we obtain

$$
\|\mathbf{v}_h - \mathbf{u}_h\|_\varepsilon \leq \frac{\beta_{\mathcal{B}_h}}{\alpha_{\mathcal{B}_h}} \|\mathbf{u} - \mathbf{v}_h\|_\varepsilon
\tag{1.2.58}
$$

and with (1.2.56) the *discrete error* depends (except for a constant factor) only on the *approximation error*:

$$
\|\mathbf{u} - \mathbf{u}_h\|_\varepsilon^2 \leq (1 + \frac{\beta_{\mathcal{B}_h}}{\alpha_{\mathcal{B}_h}}) \|\mathbf{u} - \mathbf{v}_h\|_\varepsilon^2
\tag{1.2.59}
$$

As the choice of $\mathbf{v}_h \in V_{h,0}$ is arbitrary, the same result also holds true for the infimum:

$$
\|\mathbf{u} - \mathbf{u}_h\|_\varepsilon^2 \leq (1 + \frac{\beta_{\mathcal{B}_h}}{\alpha_{\mathcal{B}_h}}) \inf_{\mathbf{v}_h \in V_{h,D}} \|\mathbf{u} - \mathbf{v}_h\|_\varepsilon^2
\tag{1.2.60}
$$

Note that the constant $C = 1 + \frac{\beta_{\mathcal{B}_h}}{\alpha_{\mathcal{B}_h}}$ only depends on the constants of the recent propositions s.t. $1 + \frac{\beta_{\mathcal{B}_h}}{\alpha_{\mathcal{B}_h}} = 1 + \sup_{\underline{x} \in \mathcal{F}_h} \left(\frac{1 + \tau_h h}{2} \right) \sqrt{1 + c_{sr}^2}$. $\square$

Now error estimates in the HDG norm are easy to get:

Proposition 1.2.8 (HDG norm estimate). *Let $\mathcal{T}_h$ be a quasi-uniform shape regular mesh, $\mathbf{u}_h \in V_h^k$ be the solution of* (1.2.9) *and $u \in H^1(\Omega) \cap H^m(\mathcal{T}_h) \subset W$, $m \geq 2$ the solution of* (1.2.1). *Then there holds the following error estimate:*

$$
\|\mathbf{u} - \mathbf{u}_h\|_\varepsilon \leq C \sqrt{\|\varepsilon\|}_\infty h^s |u|_{H^m(\mathcal{T}_h)} \qquad s = \min(k, m-1)
\tag{1.2.61}
$$

with $C \in \mathbb{R}$ only depending on the shape regularity, and the choice of the stabilization parameter τ_h

Proof. The claim follows with Proposition 1.2.6 and Lemma 1.2.7. $\square$

L^2 norm estimates

As approximation theory predicts, an approximation of h^{k+1} in the L^2-Norm is possible if we assume the problem to be sufficiently smooth. That this can actually be achieved with the proposed HDG method, we will show next with the help of the *Aubin Nitsche trick*.

Lemma 1.2.9 (Aubin Nitsche trick). *Let* $\mathbf{u}_h \in V_h^k$ *be the solution of* (1.2.9) *and* $u \in H^1(\Omega) \cap H^2(\mathcal{T}_h) \subset W$ *the solution of* (1.2.1), *both with homogeneous Dirichlet boundary conditions. Then the error converges one order faster in the L^2-norm than in the HDG norm for quasi-uniform meshes* ($h \simeq h_T \ \forall \ T \in \mathcal{T}$)

$$\|u - u_h\|_{L^2} \leq c \sqrt{\|\varepsilon\|_\infty} \ h \|\mathbf{u} - \mathbf{u}_h\|_\varepsilon \tag{1.2.62}$$

Proof. We consider the adjoint problem to (1.2.1)

$$-\mathrm{div}(\varepsilon \nabla w) = u - u_h \qquad \text{in } \Omega \tag{1.2.63}$$

with homogeneous Dirichlet boundary conditions. $u - u_h$ is in $V_D \oplus V_{D,h} \subset L^2(\Omega)$ and with a sufficiently smooth boundary $\partial\Omega$ we obtain[9] $w \in H^2(\Omega) \subset V$. As our method is also adjoint consistent (see, 1.2.37), there holds

$$\mathcal{B}_h(\mathbf{v}, \mathbf{w}) = (u - u_h, v)_\Omega \qquad \forall \ \mathbf{v} \in V_0 \tag{1.2.64}$$

Testing with $\mathbf{v} = \mathbf{u} - \mathbf{u}_h$ gives

$$\|u - u_h\|_{L^2}^2 = (u - u_h, u - u_h)_\Omega = \mathcal{B}_h(\mathbf{u} - \mathbf{u}_h, \mathbf{w}) \tag{1.2.65}$$

From (1.2.38) we know that $\mathcal{B}_h(\mathbf{u} - \mathbf{u}_h, \mathbf{v}_h) = 0$ for every $\mathbf{v}_h \in V_{h,0}$. If we choose $\mathbf{v}_h$ to be the (discontinuous) linear L^2-projection of w, $\mathbf{v}_h = \mathbf{\Pi}_{h,1} w$, there holds

$$\|u - u_h\|_{L^2}^2 = \mathcal{B}_h(\mathbf{u} - \mathbf{u}_h, \mathbf{w}) = \mathcal{B}_h(\mathbf{u} - \mathbf{u}_h, \mathbf{w} - \mathbf{\Pi}_{h,1} w) \overset{(1.2.46)}{\leq} \beta_{\mathcal{B}_h} \|\mathbf{u} - \mathbf{u}_h\|_\varepsilon \|\mathbf{w} - \mathbf{\Pi}_{h,1} w\|_\varepsilon \tag{1.2.66}$$

Now we can use standard results for the interpolation error $\|\mathbf{w} - \mathbf{\Pi}_{h,1} w\|_\varepsilon \leq C\sqrt{\|\varepsilon\|_\infty} h |w|_{H^2}$ and elliptic regularity (for sufficiently smooth boundaries) $|w|_{H^2} \leq \tilde{c}\|u - u_h\|_{L^2}$ to get

$$\|u - u_h\|_{L^2} \leq \underbrace{\beta_{\mathcal{B}_h} \tilde{c} \ C}_{=c} \sqrt{\|\varepsilon\|_\infty} h \|\mathbf{u} - \mathbf{u}_h\|_\varepsilon \tag{1.2.67}$$

$\square$

Remark 1.2.9 (nonhomogeneous Dirichlet boundary conditions):
The results also hold true for nonhomogeneous Dirichlet boundary conditions but the proof is more involved for the more general case in general $u - u_h \notin V_0$.

And finally we obtain an error estimate in the L^2-Norm:

Proposition 1.2.10 (L^2 norm estimate). *Let* $\mathcal{T}_h$ *be a quasi-uniform shape regular mesh,* $\mathbf{u}_h \in V_h^k$ *be the solution of* (1.2.9) *and* $u \in H^1(\Omega) \cap H^m(\mathcal{T}_h) \subset W$, $m \geq 2$ *the solution of* (1.2.1). *Then there holds the following error estimate:*

$$\|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)} \leq C \ \|\varepsilon\|_\infty \ h^{s+1} |u|_{H^m(\mathcal{T}_h)} \qquad s = \min(k, m-1) \tag{1.2.68}$$

with $C \in \mathbb{R}$ only depending on the shape regularity, and the choice of the stabilization parameter τ_h

Proof. The claim follows with Proposition 1.2.8 and Lemma 1.2.9. $\square$

---

[9]see for example [Evan98, 6.3]

1.2.4   Conservation property of the Hybrid DG method

One important property of most DG methods and also the proposed HDG method is the conservation of a discrete flux $\underline{\sigma}_h$. As pointed out when deriving the Hybrid DG method as a modified mixed method the conservation property appears as an explicit constraint in the variational formulation. So there holds:

Proposition 1.2.11. *The proposed Hybrid DG method is conservative (locally and globally). Furthermore we can determine a single-valued discrete flux $\underline{\sigma}_h \cdot \underline{n}$ on each facet.*

Proof. If we test the bilinearform $\mathcal{B}_h$ with $(\chi_T, 0)$, where χ_T is the characteristic function of one element, we get

$$\mathcal{B}_h(\mathbf{u}, (\chi_T, 0)) = \int_{\partial T} \varepsilon \left(-\frac{\partial u}{\partial \underline{n}} + \tau_h \, [\![\mathbf{u}]\!] \right) 1 \, d\underline{s} = \int_T f \, d\underline{x} \qquad (1.2.69)$$

With respect to the numerical flux $\underline{\sigma}_h \cdot \underline{n} := \varepsilon \left(-\frac{\partial u}{\partial \underline{n}} + \tau_h [\![\mathbf{u}]\!] \right)$ we have that the total flux which enters the element is equal to the sum of internal sources. Using this flux we have a locally conservative formulation.

To show that we don't lose anything when adding all elements together we have to show that the numerical flux $\underline{\sigma}_h \cdot \underline{n}$ is single valued on each facet, s.t.

$$\underline{\sigma}_h^+ \underline{n}^+ + \underline{\sigma}_h^- \underline{n}^- = 0 \qquad (1.2.70)$$

for all inner facets. We show this facet by facet. Let E be an inner facet of the mesh. We choose $\tilde{\mathbf{v}} = (0, v_F)$, where v_F has only support on E. Now if we plug this into the bilinearform, there holds:

$$\mathcal{B}_h(\mathbf{u}, \tilde{\mathbf{v}}) = \sum_{T \in \mathcal{T}_h} \int_{\partial T} \varepsilon \left(\frac{\partial u}{\partial \underline{n}} - \tau_h \, [\![\mathbf{u}]\!] \right) v_F \, d\underline{s} = 0 \qquad (1.2.71)$$

As v_F is chosen to have only support on E we can rewrite this as

$$\mathcal{B}_h(\mathbf{u}, \tilde{\mathbf{v}}) = \int_E \left(\underline{\sigma}_h^+ \underline{n}^+ + \underline{\sigma}_h^- \underline{n}^- \right) v_F \, d\underline{s} = 0 \qquad (1.2.72)$$

If the numerical flux $\underline{\sigma}_h \cdot \underline{n}$ is at most of order k where $v_F \in V_h^k$, equation (1.2.70) holds in the strong sense[10]. Otherwise it is still fulfilled in a weak sense. So the formulation is globally (strongly/weakly) conservative. $\qquad \square$

1.2.5   Numerical Example: Two dimensional incompressible potential flow

Problem description

We consider an irrotational incompressible flow field $\underline{v} = (v_x, v_y)$ around a circular disk. Due to irrotationality there holds $\mathrm{curl}(\underline{v}) = 0$ and so there exists a vector potential ψ with $\underline{\nabla}\psi = (v_y, -v_x)$. Furthermore there holds $\mathrm{div}(\underline{v}) = \mathrm{div}(\underline{\nabla}\psi) = -\Delta\psi = 0$ as the fluid is incompressible. We will solve for the flow potential, which is often called the *stream function* as the isolines of it coincide with the stream lines of the flow field. The geometry is a rectangular $(-1, 1) \times (-\frac{1}{2}, \frac{1}{2})$ without a disk of

---

[10]The assumption is true as long as ε is element-piecewise constant.

diameter $\frac{1}{2}$ which is removed in the middle. We prescribe Dirichlet boundaries on $\partial[(-1,1) \times (-\frac{1}{2}, \frac{1}{2})]$ and homogeneous ones on the disk (this corresponds to $\underline{\nabla}\psi \times \underline{n} = \underline{v} \cdot \underline{n} = 0$). Inhomogeneous Dirichlet boundaries are taken from the exact solution

$$\psi(x,y) = \frac{y(x^2 + y^2 - R^2)}{x^2 + y^2} \qquad R = \frac{1}{4}$$

(a) Setup of the problem: red boundaries are inhom. Dirichlet boundaries, blue boundaries are homogeneous ones

(b) Initial mesh (Level 0) and discrete solution ($p = 4$)
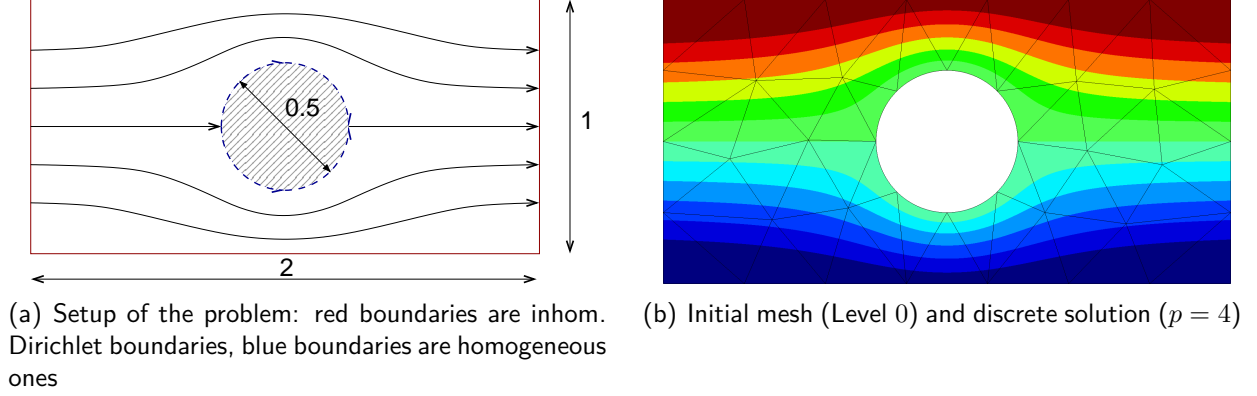
Figure 1.2.4: Setup, initial mesh and solution of Numerical Example 1.2.5

On the choice of τ_h

First we want to investigate the sensitivity of the error with respect to the stability parameter. Motivated by the estimates for the inverse inequality for simplices quoted in A.2, we choose

$$\tau_h = \alpha \frac{(p+1)(p+d)}{d} \frac{|\partial T|}{|T|} \quad, \qquad d = 2$$

As our mesh uses affine transformed as well as curved simplices we replace $|\partial T|$ by the absolute value of the jacobian determinant of the transformation of the facet and $|T|$ by the absolute value of the jacobian determinant of the transformation of the element at each integration point.

(a) error-dependency on α for $p = 1$, $L = 0$

(b) error-dependency on α for $p = 4$, $L = 0$
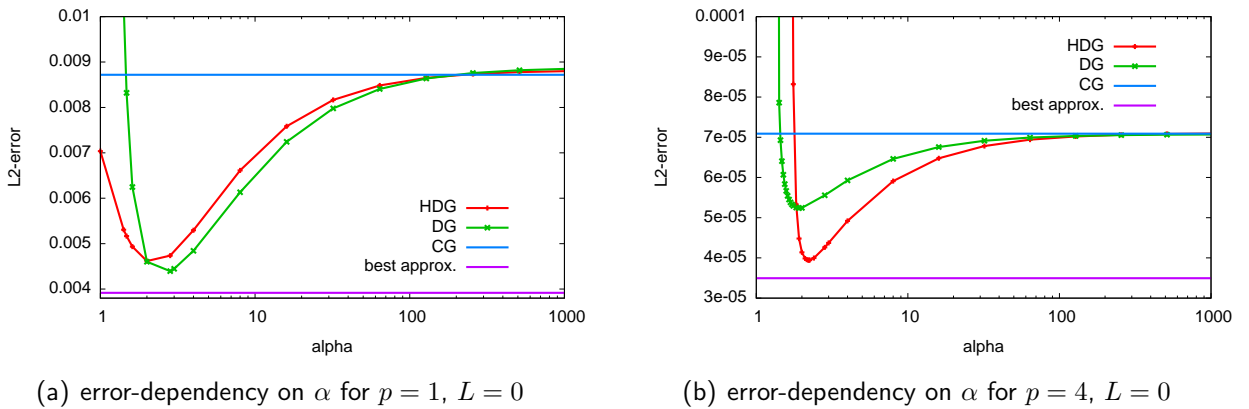
Figure 1.2.5: Sensitivity of the error on α

From the analysis we know that α has to be chosen sufficiently large for stability but not too large in order not to raise the constant of the boundedness estimate. We claimed that for well shaped meshes

this parameter is of order one. In Figure 1.2.5 the error is plotted for two choices of the polynomial order. We see that $\alpha = 2$ is a good choice and that the behaviour does not differ much between DG[11] and HDG. Furthermore the proposed scaling of τ_h motivated by affine transformed simplices works quiet well also for the curved elements. The approximation of a CG method on the same mesh with the same polynomial degree (this means a lot less degrees of freedom) is comparable with an α which has been chosen too large. That behaviour is expected as α penalizes the discontinuities and for $\alpha \to \infty$ a continuous approximation is approached.

We conclude that α is a parameter which you can choose to be 2 or a bit higher for less regular meshes, but you don't have to "optimize" the parameter for every new problem, unless you want to reduce even the least few percentages of the *discrete error*.

Notice that the solution is smooth and so the comparison to CG might not be as persuasive as in other cases. But this is not the intension of this example.

Convergence of the method

Table 1.1 confirms the L^2 norm estimates of the a priori error analysis. For $p = 1$ we get a convergence rates of 2 and for $p = 4$ the rate of 5.

meshlevel (elements)	p	unknowns DOF	$\|\mathbf{u} - \mathbf{u}_h\|_{L^2}$ error	order	p	unknowns DOF	$\|\mathbf{u} - \mathbf{u}_h\|_{L^2}$ error	order
0 (54)	1	354	4.601e-3	—	4	1 290	4.141e-5	—
1 (216)		1 356	1.516e-3	1.602		5 010	3.790e-6	3.450
2 (864)		5 304	4.028e-4	1.912		19 740	1.398e-7	4.761
3 (3 456)		20 976	1.032e-4	1.965		78 360	4.996e-9	4.806
4 (13 824)		83 424	2.605e-5	1.986		312 240	1.652e-10	4.918
5 (55 296)		332 736	6.533e-6	1.995		1 246 560	5.253e-12	4.975
6 (221 184)		1 329 024	1.636e-6	1.998				

Table 1.1: L^2-Error-Convergence of the HDG method for the potential flow problem

Comparison to Standard symmetric interior penalty DG method

We now want to present a comparison of DG and HDG. Both are expected to have with same order of convergence and for this numerical example the expectations are fulfilled. However, we now want to observe how the error decreases with respect to the total number of degrees of freedom and as a second measure with respect to the nonzero entries that are generated in the matrix. Both measures discriminate the HDG formulation, as it can be made more efficient with static condensation. Nevertheless we will see, that both measures already indicate that the HDG method without any further optimization is comparable and even cheaper than the DG method.

On the same mesh, using the same polynomial degree, the DG and the HDG formulation should have similar L^2 norm errors. Thus DG yields smaller errors for the same number of degrees of freedoms. But the gap between both methods is not dramatically large. Furthermore if we consider the second measurement, the nonzero entries in the matrix, we see in Figure 1.2.6 that the curves of DG and HDG are very close and for higher polynomial degree HDG performs even better than DG.

---

[11]We used the symmetric interior penalty DG formulation

Thus even without taking advantage of the matrix properties (which will be discussed at the end of this chapter)the HDG method is comparable or even preferable to the DG method.

Furthermore the DOF-plot shows the expected convergence rates $O(h^{p+1}) = O(N^{-\frac{p+1}{2}})$ for both methods.
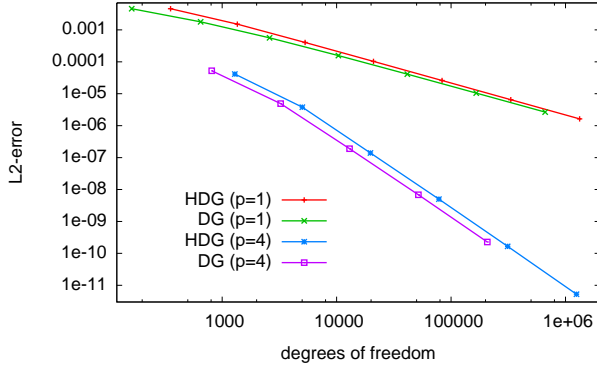
(a) error over degrees of freedom   (b) error over nonzero matrix entries
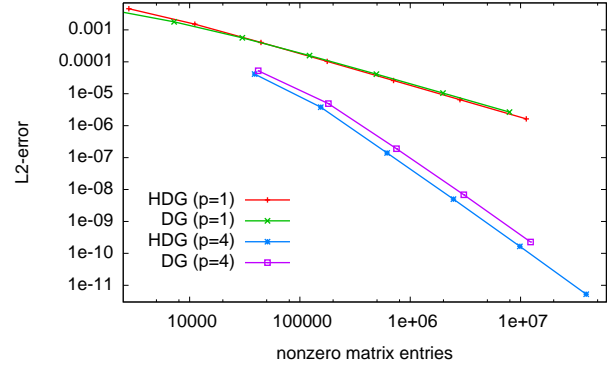
Figure 1.2.6: Possible (incomplete) measurements of the computational effort

1.3   The Hybrid Discontinuous Galerkin Method for (linear) hyperbolic problems

1.3.1   Introducing the method

In this section we consider the other limit of the convection diffusion equation, the linear transport equation, which is of hyperbolic type:

$$\begin{cases} \text{div}(\underline{b}u) & = f \qquad \text{in } \Omega \\ u & = u_D \qquad \text{on } \Gamma_{in} \end{cases} \tag{1.3.1}$$

where $\underline{b}$ has continuous normal-component b_n on element interfaces.

This time we only consider two types of boundary conditions. These are inflow and outflow boundary conditions. Boundary conditions on characteristic boundaries (these are boundaries where b_n is zero) don't influence the solution in the domain and are therefore ignored.

Linear hyperbolic problems can be solved very conveniently by Discontinuous Galerkin method. Here we use the Standard DG upwind formulation as a starting point for our Hybrid DG method. As we later on want to join the numerical scheme with the HDG formulation for the laplace operator, without losing the desired features, we have to modify the Standard DG upwind formulation a little.

The Standard DG upwind formulation for (1.3.1) is derived by partial integration on each element and choosing the upwind value for the element boundary integral:

$$\int_{\Omega} \text{div}(\underline{b}u)v \, d\underline{x} = \sum_{T \in \mathcal{T}_h} \int_T \text{div}(\underline{b}u)v \, d\underline{x} \stackrel{DG}{\approx} \sum_{T \in \mathcal{T}_h} \left\{ -\int_T \underline{b}u \underline{\nabla} v \, d\underline{x} + \int_{\partial T} b_n u^{DG} \, v \, d\underline{s} \right\} =: \mathcal{C}_h^{DG}(u,v)$$

where $u^{DG} = \begin{cases} u & b_n > 0 \\ u_{nb} & b_n < 0 \end{cases}$ with u_{nb} the discrete trace on the neighbour element. This means that we always choose the value which lies in the direction from where the convection originates.

Now we modify the formulation such that the access to the neighbour element functions is replaced by an access to the facet functions which we already introduced in section 1.2. So on outflow parts of the element boundary we choose the element value $u^{up} = u$ (see Fig. 1.3.1(a)). But on inflow parts of the element boundary we choose, in contrast to Standard DG methods, the facet value $u^{up} = u_F$ (see Fig. 1.3.1(b)). This has the advantage that we preserve the property of the elliptic hdg formulation, where element degrees of freedom from different elements don't couple directly.

$$\sum_{T \in \mathcal{T}_h} \int_T \text{div}(\underline{b}u)v \, d\underline{x} \stackrel{HDG}{\approx} \sum_{T \in \mathcal{T}_h} \left\{ -\int_T \underline{b}u \underline{\nabla} v \, d\underline{x} + \int_{\partial T} b_n u^{up} \, v \, d\underline{s} \right\} \text{ with } u^{up} = \begin{cases} u & b_n > 0 \\ u_F & b_n < 0 \end{cases}$$

$$\tag{1.3.2}$$

(a) at the outflow of an element (green, surrounded by dashed line) the element value (red, hatched) is taken. $u^{up} = u$

(b) at the inflow of an element (green, surrounded by dashed line) the facet value (red, bold) is taken. $u^{up} = u_F$
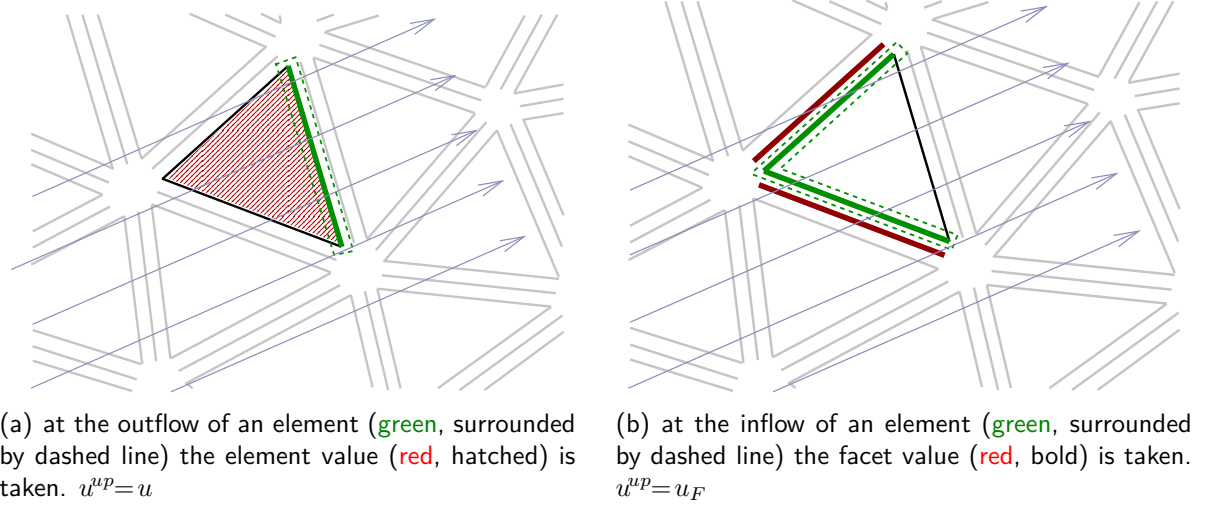
Figure 1.3.1: Choice for the upwind value for the HDG formulation for one element

It is obvious that now the unknowns of different elements don't even couple at all, because facet unknowns just couple with one neighbouring element, the downwind element, and not with upwind element. We overcome this issue by adding a constraint which glues the facet values on the trace of the upwind element (in a weak sense):

$$\sum_{T \in \mathcal{T}_h} \int_{\partial T_{out}} b_n (u_F - u)\, v_F \, d\underline{s} = 0 \qquad \text{with} \qquad \partial T_{out} := \{\underline{x} \in \partial T, b_n(\underline{x}) > 0\} \qquad (1.3.3)$$

The "stabilization" is illustrated in Fig. 1.3.2 and results in $u_F = u^{up}$
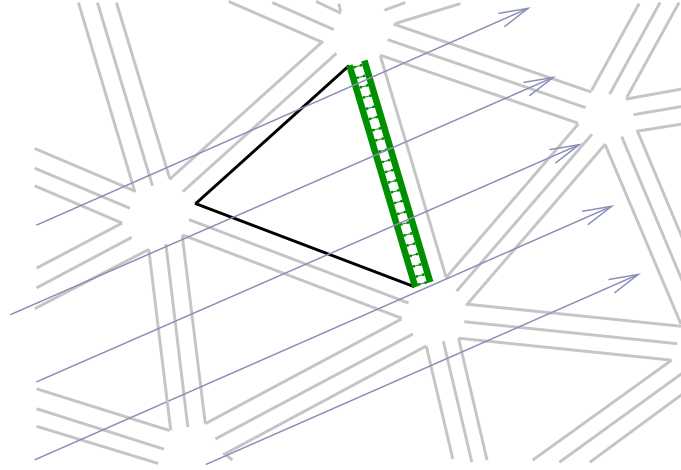
Figure 1.3.2: at the outflow the facet value and the upwind element value are glued together

Adding up both equations (1.3.3) and (1.3.2) we end up with

$$\mathcal{C}_h(\mathbf{u}, \mathbf{v}) := \sum_{T \in \mathcal{T}_h} \left\{ - \int_T \underline{b} u \underline{\nabla} v \, d\underline{x} + \int_{\partial T} b_n u^{up}\, v \, d\underline{s} + \int_{\partial T_{out}} b_n (u_F - u)\, v_F \, d\underline{s} \right\} = \int_\Omega f\, v \, d\underline{x} \quad (1.3.4a)$$

You can rewrite this formulation also in the following form which is sometimes more convenient for the analysis by doing partial integration once more and noting that the first boundary term vanishes on

∂T_{out} as u^{up} is exactly u there and b_n is positive on ∂T_{out} and negative on $\partial T_{in} := \{\underline{x} \in \partial T, b_n(\underline{x}) < 0\}$

$$\mathcal{C}_h(\mathbf{u}, \mathbf{v}) = \sum_{T \in \mathcal{T}_h} \left\{ \int_T \text{div}(\underline{b}u)v \, d\underline{x} + \int_{\partial T_{in}} |b_n|(u - u_F) \, v \, d\underline{s} + \int_{\partial T_{out}} |b_n|(u_F - u) \, v_F \, d\underline{s} \right\} \qquad (1.3.4b)$$

In this form you especially see that the chosen sign for the stabilization term results in positive entries on the matrix of the discretized linear system.

A third way of presenting the bilinearform, which will be helpful when looking at the discrete adjoint of $\mathcal{C}_h$, can be achieved if we move $u_F v_F$ from ∂T_{out} to ∂T_{in}, or, more generally spoken (to consider the boundary facet in a correct way), we first substract $\int_{\partial T_{out}} |b_n| u_F v_F \, d\underline{s}$ and add $\int_{\partial T_{in}} |b_n| u_F v_F \, d\underline{s}$ on each element $(*)$. For inner facets both terms cancel out and the result is a simple shift from the in- to the outflow side. On boundary facets we have to add additional boundary integrals for balance. Starting from (1.3.4b), after integration by parts and the described procedure, we can formulate $\mathcal{C}_h$ also in the following way:

$$\begin{aligned}
\mathcal{C}_h(\mathbf{u}, \mathbf{v}) &= \sum_{T \in \mathcal{T}_h} \left\{ -\int_T \underline{b}u \nabla v \, d\underline{x} + \int_{\partial T_{in}} |b_n|(-u_F) \, v \, d\underline{s} + \int_{\partial T_{out}} |b_n|(u \, v + (u_F - u) \, v_F) \, d\underline{s} \right\} \\
&\overset{(*)}{=} \sum_{T \in \mathcal{T}_h} \left\{ -\int_T \underline{b}u \nabla v \, d\underline{x} + \int_{\partial T} b_n u^{up} \, [\![v]\!] \, d\underline{s} + \int_{\partial T \cap \Gamma} b_n u_F v_F \, d\underline{s} \right\}
\end{aligned}$$
$$(1.3.4c)$$

All three formulations are algebraically equivalent with Standard DG upwind, but also exhibit the same nice properties of HDG methods that we already saw for the elliptic case and will also be elaborated later on:

- element-wise assembly is possible

- element unknowns of different elements don't couple directly

- inner degrees of freedom can be eliminated, i.e. the local problems are uniquely solvable

The discrete problem reads:

$$\text{Find } \mathbf{u} \in V_D, \text{ such that for all } \mathbf{v} \in V_{h,0} :$$
$$\mathcal{C}_h(\mathbf{u}, \mathbf{v}) = \langle f, v \rangle = \int_\Omega f \, v \, d\underline{x} \qquad (1.3.5)$$

Remark 1.3.1 (Discontinuities):
Equation (1.3.1) could have discontinuous solutions. Those can be captured exactly by the method as long as the mesh is aligned to the discontinuity. That's because for a mesh aligned to the discontinuity we have $b_n = 0$ and so the problems on each side of the discontinuity decouple also for the numerical scheme. Notice that as we assume that at least b_n is continuous and so is zero on both sides of a discontinuity, the unknowns and testfunctions u_F and v_F disappear completely from the formulation. In order to make the matrix of the discretized bilinearform regular again, we could add integrals of the form $\int_{\partial T^*} |\underline{b}| [\![\mathbf{u}]\!] v_F \, d\underline{s}$ on element boundaries $\partial T^* = \{\underline{x} \in \partial T, b_n(\underline{x}) = 0\}$ which are aligned to characteristics. Those integrals wouldn't influence the solution on the neighbouring domains, but give convenient defining equations for u_F which for equal order interpolation and single-valued $\underline{b}$ on each facet give $u_F = \{\![u]\!\}$. Notice that this choice for u_F is arbitrary as there is no meaningful value to describe at the discontinuity.

1.3.2 A priori error analysis

After the method is introduced, we want to carry out an a priori error analysis for the HDG upwind formulation. Although we already pointed out that the proposed formulation is equivalent to the Standard DG method, the analysis will be carried out completely and in a compatible way that allows for unification with the discretization for the elliptic operator.

Basic assumptions

We assume $\mathrm{div}(\underline{b}) = 0$, i.e. $\underline{b}$ a solenoidal vector. To pay tribute to the hyperbolic nature of (1.3.1) we replace the assumption $u \in W$ of section 1.2 by $u \in W^{conv} = \{v \in L^2(\Omega); \underline{b} \cdot \underline{\nabla} v \in L^2(\Omega)\}$ instead. That means that we, in contrast to the assumption $u \in W$, allow for discontinuities along streamlines. For ease of notation we will also call it W during the whole analysis of the pure hyperbolic problem. Further we also use $V := W \times L^2(\mathcal{F}_h)$. The discrete spaces are not exchanged. Similar to the notation of the pure elliptic case V_D incorporates Dirichlet boundary conditions on the boundary. It only makes sense to prescribe Dirichlet boundary conditions on $\Gamma_{in} = \{x \in \Gamma, b_n < 0\}$ or $\Gamma_{out} = \{x \in \Gamma, b_n > 0\}$. We assume $\Gamma_D = \Gamma_{in}$ to follow the physical causality of the equation.

We further assume that discontinuities are not aligned to the mesh, meaning that $b_n = 0$ doesn't hold on a single facet. As mentioned in Remark 1.3.1 we could also consider this case easily, but to avoid a blow-up in notation we forego that.

In this section we want to devide the bilinearform into two parts

$$\mathcal{C}_h(\cdot, \cdot) = \mathcal{C}_h^{el}(\cdot, \cdot) + \mathcal{C}_h^{fac}(\cdot, \cdot) \tag{1.3.6}$$

where $\mathcal{C}_h^{el}(\cdot, \cdot)$ includes all integrals on the volume of each element and $\mathcal{C}_h^{fac}(\cdot, \cdot)$ includes all element boundary (facet) integrals.

Remember also that we assumed $\mathrm{div}(\underline{b}) = 0$.

1.3.2.1 Consistency

We now expand the true solution of (1.3.1) by the facet values $u_F(u)$ and with the regularity along streamlines of V^{conv} we are able to define the upwind trace of the true solution $u_F(\underline{x}) = \lim_{\epsilon \to 0^+} u(\underline{x} - \epsilon \underline{b})$ on each facet as long as $b_n \neq 0$. [12] We use the representation (1.3.4b) to show consistency. Let $\mathbf{u} \in V_D$ be the solution of (1.3.1), then for all $v \in V$

$$\mathcal{C}_h(\mathbf{u}, \mathbf{v}) = \sum_{T \in \mathcal{T}_h} \left\{ \int_T \underbrace{\mathrm{div}(\underline{b}u)}_{=f} v \, d\underline{x} + \int_{\partial T_{in}} \underbrace{|b_n|(u - u_F)}_{=0} v \, d\underline{s} + \int_{\partial T_{out}} \underbrace{|b_n|(u_F - u)}_{=0} v_F \, d\underline{s} \right\} = \int_\Omega f \, v \, d\underline{x} \tag{1.3.7}$$

holds. We again directly obtain Galerkin orthogonality from that.

Proposition 1.3.1 (Galerkin Orthogonality)**.** *Let $\mathbf{u}_h \in V_{h,D}$ be the solution of (1.3.5) and u be the solution of (1.3.1). Then there holds*

$$\mathcal{C}_h(\mathbf{u}_h - \mathbf{u}, \mathbf{v}_h) = 0 \qquad \forall \, \mathbf{v}_h \in V_h \tag{1.3.8}$$

---

[12]In the case $b_n = 0$ the definition is irrelevant because of the decoupling of the values on the neighbouring elements

Proof. Using (1.3.7) we get: For all $\mathbf{v}_h \in V_h$

$$\mathcal{C}_h(\mathbf{u} - \mathbf{u}_h, \mathbf{v}_h) = \int_\Omega f \; v_h \, d\underline{x} - \int_\Omega f \; v_h \, d\underline{x} = 0 \tag{1.3.9}$$

$\square$

We are also interested in the adjoint consistency of our formulation. Therefore we pose the adjoint problem to (1.3.1):

$$\tilde{\underline{b}} \cdot \underline{\nabla} u = f \quad \text{in } \Omega ; \quad u = 0 \quad \text{on } \Gamma_{out} = \tilde{\Gamma}_{in}; \text{ with } \tilde{\underline{b}} = -\underline{b} \tag{1.3.10}$$

Note that the boundaries without tilde still refer to the original problem. The outflow boundary of the original problem Γ_{out} is the inflow boundary $\tilde{\Gamma}_{in}$ of the adjoint problem. In the case $\mathrm{div}(\underline{b}) = 0$ the adjoint problem is the same as the original one except that the convective velocity is $\tilde{\underline{b}} = -\underline{b}$ now and the role of in- and outflow boundary are exchanged. After discretization (using representation (1.3.4c)) we conclude that the discrete adjoint $\mathcal{C}_h(\mathbf{v}, \mathbf{u})$ is consistent with the adjoint problem (1.3.10).

$$\mathcal{C}_h(\mathbf{v}, \mathbf{u}) = \sum_{T \in \mathcal{T}_h} \left\{ \int_T \underbrace{(\tilde{\underline{b}} \cdot \underline{\nabla} u)}_{f} v \, d\underline{x} - \int_{\partial T} \tilde{b}_n u^{up} \underbrace{[\![\mathbf{v}]\!]}_{=0} \, d\underline{s} - \int_{\partial T \cap \Gamma} \tilde{b}_n \underbrace{v_F u_F}_{=0 \text{ on } \Gamma} \, d\underline{s} \right\} \tag{1.3.11}$$

where $v_F u_F = 0$ on Γ holds since u_F is zero on the inflow boundary of the adjoint problem which is $\tilde{\Gamma}_{in} = \Gamma_{out}$ and v_F is assumed to be zero on the outflow boundary of the adjoint problem which is $\Gamma_{in} = \tilde{\Gamma}_{out}$. So we directly get

$$\mathcal{C}_h(\mathbf{v}, \mathbf{u}) = \int_\Omega fv \, d\underline{x} \quad \forall \, \mathbf{v} \in V_0 \tag{1.3.12}$$

for the solution of the adjoint problem.

1.3.2.2   Stability

In this paragraph we assume $u = 0$ on $\Gamma_{in} = \Gamma_D$ which is reasonable as we will need the stability result for $\mathbf{u}_h u - \mathbf{v}_h$ with $\mathbf{u}_h, \mathbf{u}_h$ in $V_{h,D}$ and thus $\mathbf{u}_h u - \mathbf{v}_h \in V_{h,0}$. Here we follow the line of arguments of [ES09] to show stability of the (hybrid) DG method in a way which allows for generalization with the elliptic formulation later on (see section 1.2.3.2).
After introducing the *convection HDG norm*

$$\|\mathbf{u}\|_{C,*}^2 := \sum_{T \in \mathcal{T}_h} \left\{ \int_T \frac{h}{|\underline{b}|} (\underline{b} \cdot \underline{\nabla} u)^2 \, d\underline{x} + \int_{\partial T} |b_n| [\![\mathbf{u}]\!]^2 \, d\underline{s} + \int_{\partial T \cap \Gamma_{out}} |b_n| u_F^2 \, d\underline{s} \right\} = \sum_{T \in \mathcal{T}_h} (\|\mathbf{u}\|_{C,*}^T)^2 \tag{1.3.13}$$

we claim the stability statement, known as inf-sup stability

Proposition 1.3.2 (inf-sup stability). *Let u be the exact solution of (1.3.1). Then, for a constant $\alpha_{\mathcal{C}_h} \in \mathbb{R}$ independent of the meshsize h the following is true:*

$$\sup_{\mathbf{v} \in V_{h,0}} \frac{\mathcal{C}_h(\mathbf{u}, \mathbf{v})}{\|\mathbf{v}\|_{C,*}} \geq \alpha_{\mathcal{C}_h} \|\mathbf{u}\|_{C,*} \quad \forall \, \mathbf{u} \in V_{h,0} \tag{1.3.14}$$

Furthermore, for every $\mathbf{u} \in V_{h,0}$ we can find a function $\tilde{\mathbf{v}} \in V_{h,0}$, s.t.

$$\mathcal{C}_h(\mathbf{u}, \tilde{\mathbf{v}}) \geq \alpha_{\mathcal{C}_h} \|\mathbf{u}\|_{C,*} \|\tilde{\mathbf{v}}\|_{C,*} \tag{1.3.15}$$

Proof. We will show that the *inf-sup-condition* holds with a particular choice for $\mathbf{v} \in V_{h,0}$:

$$\mathbf{v} = (v, v_F) = c_1(u, u_F) + c_2\left(\frac{h}{|\underline{b}|}\underline{b} \cdot \underline{\nabla}u, 0\right) \qquad (1.3.16)$$

With this approach we will show the second part of the Proposition which implies the first one.
First we want to present the terms coming up for the choices $\mathbf{v} = \mathbf{u}$ and $\mathbf{v} = (\frac{h}{|\underline{b}|}\underline{b} \cdot \underline{\nabla}u, 0)$ separately and afterwards pick one possible linear combination for $\tilde{\mathbf{v}}$.
Notice that as we have assumed that $\mathrm{div}(\underline{b}) = 0$ there holds:

$$\int_T \underline{b} \cdot \underline{\nabla}u \; u \, d\underline{x} = \int_T \mathrm{div}(\underline{b}u) \; u \, d\underline{x} \overset{p.I.}{=} -\int_T \underline{b} \cdot \underline{\nabla}u \; u \, d\underline{x} + \int_{\partial T} b_n u^2 \, d\underline{s}$$

$$\Rightarrow \int_T \mathrm{div}(\underline{b}u) \; u \, d\underline{x} = \frac{1}{2}\int_{\partial T} b_n u^2 \, d\underline{s} \qquad (1.3.17)$$

Now we use the representation (1.3.4b) for our bilinearform and test with $\mathbf{v} = \mathbf{u}$:

$$
\begin{aligned}
\mathcal{C}_h(\mathbf{u}, \mathbf{u}) &= \sum_{T \in \mathcal{T}_h}\left\{\int_T \mathrm{div}(\underline{b}u)u \, d\underline{x} + \int_{\partial T_{in}}|b_n|(u - u_F)\, u \, d\underline{s} + \int_{\partial T_{out}}|b_n|(u_F - u)\, u_F \, d\underline{s}\right\} \\
&\overset{(1.3.17)}{=} \sum_{T \in \mathcal{T}_h}\left\{\frac{1}{2}\int_{\partial T} b_n u^2 \, d\underline{s} + \int_{\partial T_{in}}|b_n|(u - u_F)\, u \, d\underline{s} + \int_{\partial T_{out}}|b_n|(u_F - u)\, u_F \, d\underline{s}\right\} \\
&= \sum_{T \in \mathcal{T}_h}\left\{\frac{1}{2}\int_{\partial T_{in}}|b_n|(u^2 - 2\,u\,u_F)\, d\underline{s} + \frac{1}{2}\int_{\partial T_{out}}|b_n|(u^2 - 2u\,u_F + 2u_F^2)\, d\underline{s}\right\} \\
&\overset{(*)}{=} \sum_{T \in \mathcal{T}_h}\left\{\frac{1}{2}\int_{\partial T}|b_n|[\![\mathbf{u}]\!]^2 \, d\underline{s} - \frac{1}{2}\int_{\partial T \cap \Gamma_{in}}|b_n|\underbrace{u_F^2}_{=0} \, d\underline{s} + \frac{1}{2}\int_{\partial T \cap \Gamma_{out}}|b_n|u_F^2 \, d\underline{s}\right\} \\
&= \frac{1}{2}\sum_{T \in \mathcal{T}_h}\left\{\int_{\partial T}|b_n|[\![\mathbf{u}]\!]^2 \, d\underline{s} + \int_{\partial T \cap \Gamma_{out}}|b_n|u_F^2 \, d\underline{s}\right\}
\end{aligned}
$$
$$(1.3.18)$$

In $(*)$ we moved one half of u_F^2 from ∂T_{out} to ∂T_{in}, or more generally spoken (to consider the boundary facets in a correct way) we first substract $\frac{1}{2}\int_{\partial T_{out}}|b_n|u_F^2 \, d\underline{s}$ and add $\frac{1}{2}\int_{\partial T_{in}}|b_n|u_F^2 \, d\underline{s}$ on each element. For inner facets both terms cancel out and the result is a simple shift from the in- to the outflow side. On boundary facets we have to add additional boundary integrals for balance.
So we see that we can use this for the second part of the norm $\|\cdot\|_{C,*}$. The volume part of that norm will be account for with $\mathbf{v} = \mathbf{v}^* = (\frac{h}{|\underline{b}|}\underline{b} \cdot \underline{\nabla}u, 0)$. Notice that as $v_F^* = 0$ there holds $\mathbf{v}^* \in V_{h,0}$:

$$
\begin{aligned}
\mathcal{C}_h(\mathbf{u}, \mathbf{v}^*) &= \sum_{T \in \mathcal{T}_h}\left\{\int_T \frac{h}{|\underline{b}|}\, \mathrm{div}(\underline{b}u)\underline{b} \cdot \underline{\nabla}u \, d\underline{x} + \int_{\partial T_{in}}\frac{h}{|\underline{b}|}\,|b_n|\,[\![\mathbf{u}]\!]\,\underline{b} \cdot \underline{\nabla}u \, d\underline{s}\right\} \\
&\geq \sum_{T \in \mathcal{T}_h}\left\{\int_T \frac{h}{|\underline{b}|}\,(\underline{b} \cdot \underline{\nabla}u)^2 \, d\underline{x} - \int_{\partial T_{in}}\frac{h}{|\underline{b}|}\,|b_n|\,|[\![\mathbf{u}]\!]|\,|\underline{b} \cdot \underline{\nabla}u|\, d\underline{s}\right\} \\
&\overset{(A.2.2)}{\geq} \sum_{T \in \mathcal{T}_h}\left\{\int_T \frac{h}{|\underline{b}|}\,(\underline{b} \cdot \underline{\nabla}u)^2 \, d\underline{x} - \alpha\int_{\partial T_{in}}|b_n|\,[\![\mathbf{u}]\!]^2 \, d\underline{s} - \frac{1}{\alpha}\int_{\partial T_{in}}\frac{h^2}{|\underline{b}|^2}\,|b_n|\,(\underline{b} \cdot \underline{\nabla}u)^2 \, d\underline{s}\right\} \\
&\overset{|b_n|<|\underline{b}|}{\geq} \sum_{T \in \mathcal{T}_h}\left\{\int_T \frac{h}{|\underline{b}|}\,(\underline{b} \cdot \underline{\nabla}u)^2 \, d\underline{x} - \alpha\int_{\partial T_{in}}|b_n|\,[\![\mathbf{u}]\!]^2 \, d\underline{s} - \frac{1}{\alpha}\int_{\partial T_{in}}\frac{h^2}{|\underline{b}|}\,(\underline{b} \cdot \underline{\nabla}u)^2 \, d\underline{s}\right\} \\
&\overset{(A.2.3a)}{\geq} \sum_{T \in \mathcal{T}_h}\left\{\int_T \frac{h}{|\underline{b}|}\,(\underline{b} \cdot \underline{\nabla}u)^2 \, d\underline{x} - \alpha\int_{\partial T_{in}}|b_n|\,[\![\mathbf{u}]\!]^2 \, d\underline{s} - \frac{c_T^2}{\alpha}\int_T \frac{h}{|\underline{b}|}\,(\underline{b} \cdot \underline{\nabla}u)^2 \, d\underline{s}\right\} \\
&\geq \sum_{T \in \mathcal{T}_h}\left\{\left(1 - \frac{c_T^2}{\alpha}\right)\int_T \frac{h}{|\underline{b}|}\,(\underline{b} \cdot \underline{\nabla}u)^2 \, d\underline{x} - \alpha\int_{\partial T_{in}}|b_n|\,[\![\mathbf{u}]\!]^2 \, d\underline{s}\right\}
\end{aligned}
$$
$$(1.3.19)$$

Set α to $2c_{sr}^2$ s.t. $(1 - \frac{c_T^2}{\alpha})$ is greater $\frac{1}{2}$ and add sufficiently much of $\mathbf{u}$ to ensure that the total boundary contribution has a positive factor greater $\frac{1}{2}$. This is achieved with the choice $\tilde{\mathbf{v}} = (2c_{sr}^2 + 1)\mathbf{u} + \mathbf{v}^*$. The constant c_{sr} is the maximum on the element local constants $c_T \in \mathbb{R}$ which only depend on the shape regularity of the mesh. With this linear combination we get:

$$\mathcal{C}_h(\mathbf{u}, \tilde{\mathbf{v}}) \;\geq\; \tfrac{1}{2}\|\mathbf{u}\|_{C,*}^2 \tag{1.3.20}$$

Now we additionally need

$$\|\tilde{\mathbf{v}}\|_{C,*} \leq \frac{1}{2\alpha_{\mathcal{C}_h}}\|\mathbf{u}\|_{C,*} \tag{1.3.21}$$

From scaling arguments, i.e. arguments that bound the norm of gradients of polynomials by the norm of the polynomial itself and an appropriate scaling $\frac{1}{h}$, you can bound $\|\mathbf{v}^*\|_{C,*}$ by $\tilde{c}\|\mathbf{u}\|_{C,*}$. Then with the triangle inequality there holds

$$
\begin{aligned}
\|\tilde{\mathbf{v}}\|_{C,*}^2 &\leq (2c_{sr}^2 + 1)^2\|\mathbf{u}\|_{C,*}^2 + \|\mathbf{v}^*\|_{C,*}^2 \\
&\leq \underbrace{\left((2c_{sr}^2 + 1)^2 + \tilde{c}^2\right)}_{=:(\frac{1}{2\alpha_{\mathcal{C}_h}})^2}\|\mathbf{u}\|_{C,*}^2
\end{aligned}
\tag{1.3.22}
$$

And finally

$$\mathcal{C}_h(\mathbf{u}, \tilde{\mathbf{v}}) \;\geq\; \alpha_{\mathcal{C}_h}\|\mathbf{u}\|_{C,*}\|\tilde{\mathbf{v}}\|_{C,*} \tag{1.3.23}$$

Notice that $\beta \in \mathbb{R}$ only depends on fixed constants and on the shape regularity of the mesh. $\qquad\square$

1.3.2.3   Boundedness

Now we want to show boundedness of the bilinearform $\mathcal{C}_h$. As we have seen in chapter 1.2 there is not only one possible choice for an HDG norm and which one might be the more natural one may differ from boundedness- to coercivity-estimates. When dealing with the boundedness it is convenient to work with a *second convection HDG norm* $\|\cdot\|_C$:

$$\|\mathbf{u}\|_C^2 := \sum_{T \in \mathcal{T}_h}\left\{\int_T \frac{|b|}{h}u^2\, d\underline{x} + \int_{\partial T}|b_n|(u^{up})^2\, d\underline{s} + \int_{\partial T \cap \Gamma_{out}}|b_n|u_F^2\, d\underline{s}\right\} = \sum_{T \in \mathcal{T}_h}(\|\mathbf{u}\|_C^T)^2 \tag{1.3.24}$$

We can easily show with scaling arguments that for all *discrete* functions $\mathbf{u} \in V_h$

$$\|\mathbf{u}\|_{C,*}^2 \leq C\|\mathbf{u}\|_C^2 \tag{1.3.25}$$

holds.

In the elliptic case we also introduced two different norms but there it would have been sufficient to work with the latest one only. In the hyperbolic case, here, we have to work with both norms at the same time. That's why both norms are involved in the following Proposition:

Proposition 1.3.3. *For all* $\mathbf{u} \in V$ *and* $\mathbf{v} \in V_0$ *there holds:*

$$|\mathcal{C}_h(\mathbf{u}, \mathbf{v})| \leq \|\mathbf{u}\|_C\ \|\mathbf{v}\|_{C,*}$$

Proof. We want to bound $\mathcal{C}_h(\mathbf{u}, \mathbf{v})$ for $\mathbf{u} \in V$ and $\mathbf{v} \in V_0$. Therefore we start with $\mathcal{C}_h^{el}$ and $\mathcal{C}_h^{fac}$ of the representation (1.3.4c):

$$
\begin{aligned}
|\mathcal{C}_h^{el}(\mathbf{u}, \mathbf{v})| &= \left| \sum_{T \in \mathcal{T}_h} - \int_T \underline{b} u \underline{\nabla} v \, d\underline{x} \right| \\
&\leq \sum_{T \in \mathcal{T}_h} \int_T \left| |\underline{b}|^{\frac{1}{2}} h^{-\frac{1}{2}} u \right| \left| |\underline{b}|^{-\frac{1}{2}} h^{\frac{1}{2}} (\underline{b} \cdot \underline{\nabla} v) \right| \, d\underline{x} \\
&\overset{C.S}{\leq} \sum_{T \in \mathcal{T}_h} \sqrt{\int_T |\underline{b}| h^{-1} u^2 \, d\underline{x}} \sqrt{\int_T |\underline{b}|^{-1} h (\underline{b} \cdot \underline{\nabla} v)^2 \, d\underline{x}}
\end{aligned}
$$

(1.3.26)

$$
\begin{aligned}
|\mathcal{C}_h^{fac}(\mathbf{u}, \mathbf{v})| &\overset{(1.3.4c)}{=} \left| \sum_{T \in \mathcal{T}_h} \left\{ \int_{\partial T} b_n u^{up} \llbracket \mathbf{v} \rrbracket \, d\underline{s} - \int_{\partial T \cap \Gamma_{in}} |b_n| u_F \underbrace{v_F}_{=0} \, d\underline{s} + \int_{\partial T \cap \Gamma_{out}} |b_n| u_F v_F \, d\underline{s} \right\} \right| \\
&\overset{C.S}{\leq} \sum_{T \in \mathcal{T}_h} \left\{ \sqrt{\int_{\partial T} |b_n| (u^{up})^2 \, d\underline{s}} \sqrt{\int_{\partial T} |b_n| \llbracket \mathbf{v} \rrbracket^2 \, d\underline{s}} + \sqrt{\int_{\partial T \cap \Gamma_{out}} |b_n| u_F^2 \, d\underline{s}} \sqrt{\int_{\partial T \cap \Gamma_{out}} |b_n| v_F^2 \, d\underline{s}} \right\}
\end{aligned}
$$

(1.3.27)

With these preparations we can now finish the proof of boundedness:

$$
\begin{aligned}
|\mathcal{C}_h(\mathbf{u}, \mathbf{v})| &\leq |\mathcal{C}_h^{el}(\mathbf{u}, \mathbf{v})| + |\mathcal{C}_h^{fac}(\mathbf{u}, \mathbf{v})| \\
&\overset{C.S.}{\leq} \sum_{T \in \mathcal{T}_h} \vertiii{\mathbf{u}}_C^T \cdot \vertiii{\mathbf{v}}_{C,*}^T \\
&\leq \vertiii{\mathbf{u}}_C \cdot \vertiii{\mathbf{v}}_{C,*}
\end{aligned}
$$

(1.3.28)

$\square$

1.3.2.4 Approximation

As the *HDG norms* of the convection problem scale differently from those of the elliptic problem, we repeat and adapt the basic steps of 1.2.3.3. We will again assume some $H^m(\mathcal{T}_h)$-regularity which is certainly to much to ask for in most cases. A more appropriate space would be the Hilbert space with L^2-functions with weak directional derivatives up to mth order. Then for piecewise constant $\underline{b}$ an adapted version of the *Bramble Hilbert Lemma* could be stated[13]. The remainder would still be the same.

Later on, when looking at the convection diffusion problem for $\varepsilon > 0$ the full $H^m(\mathcal{T}_h)$-regularity assumption is again fulfilled for sufficiently smooth boundaries, but, due to nearly discontinuities (e.g. boundary layers), the upcoming constants (e.g. the H^m-semi norm) in the following estimates may degenerate.

We will stick with the more convenient although less appropriate $H^m(\mathcal{T}_h)$-regularity assumption and use another application of the original Bramble Hilbert Lemma.

Lemma 1.3.4. *For $u \in H^m(T)$, $m = k + 1$ and the L^2-projector $\mathbf{\Pi}_{h,k} : H^m(T) \to \mathcal{P}^k(T) \times F_h^k(T)$ of* (1.2.47), *there holds*

$$
\vertiii{(\mathbf{id}_h - \mathbf{\Pi}_{h,k})u}_C^T \leq C \sqrt{\|\underline{b}\|}_\infty |u|_{H^m(T)} \qquad \forall \, u \in H^m(T)
$$

(1.3.29)

---

[13] in this case we need piecewise constant $\underline{b}$, s.t. the mth order seminorm has the kernel $\mathcal{P}^{m-1}$

Proof. The proof goes analogously to the proof of Lemma 1.2.5. The scalar product is replaced by a scalar product which fulfills $(\mathbf{u}, \mathbf{u})_U = \|\mathbf{u}\|_U^2 \geq \left(\|\sqrt{\frac{1}{|b|}}\, \mathbf{u}\|_C^T \right)^2 + R(\mathbf{u}, \mathbf{v})$ with $R(\cdot, \cdot)$ some symmetric nonnegative bilinearform. Again boundedness of the operator $L = \mathbf{id}_h - \mathbf{\Pi}_{h,k}$ is easy to get and so all requirements for the *Bramble Hilbert Lemma* are fulfilled again and we get

$$\|(\mathbf{id}_h - \mathbf{\Pi}_{h,k})u\|_C^T \leq \|\sqrt{|b|}\, Lu\|_U \leq C \sqrt{\|b\|}_\infty\, |u|_{H^m(T)} \qquad \forall\, u \in H^m(T) \tag{1.3.30}$$

$\square$

Now we will look at the global interpolation error to bound the *approximation error* $\inf_{\mathbf{v}_h \in V_h} \|\mathbf{u} - \mathbf{v}_h\|_C$. Therefore we will transform all element contributions on a reference element. There we will use Lemma (1.3.4) and after transforming back to physical domain, we will get the right scaling of the interpolation error.

Proposition 1.3.5. *On a shape regular mesh $\mathcal{T}_h$ which consists of affine transformed elements and a function $u \in H^1(\Omega) \cap H^m(\mathcal{T}_h)$, $m = k + 1 \geq 2$ there holds the following approximation result for V_h^k as introduced in (1.2.35c)*

$$\inf_{\mathbf{v}_h \in V_h^k} \|\mathbf{u} - \mathbf{v}_h\|_C^2 \leq C \sum_{T \in \mathcal{T}_h} \|b\|_\infty h_T^{2k+1} |u|_{H^m(T)}^2 \tag{1.3.31}$$

with C independent of the mesh size. For quasi-uniform meshes the direct conclusion of it is

$$\inf_{\mathbf{v}_h \in V_h^k} \|\mathbf{u} - \mathbf{v}_h\|_C \leq C \sqrt{\|b\|_\infty}\, h^{k+\frac{1}{2}} |u|_{H^m(\mathcal{T}_h)} \tag{1.3.32}$$

Proof. Again the proof is similar to the one of Proposition 1.2.6 except for the scaling of the norm which is different to the *diffusion HDG norms* of section 1.2.3. On each element we now get

$$\left(\|(\mathbf{id}_h^T - \mathbf{\Pi}_{h,k}^T)u\|_C^T \right)^2 \preceq h_T^d h_T^{-1} \left(\|(\mathbf{id}_h^T - \mathbf{\Pi}_{h,k}^T)u \circ F_T\|_C^{\hat{T}} \right)^2 \tag{1.3.33}$$

where the scaling of the square of the norm is h_T^{d-1}, which will be explained in the following table

part of $\|\cdot\|_C^2$	scaling factor	integral trafo.	diff. op. scaling	overall scaling
$\dfrac{1}{h} \displaystyle\int_T |b| u^2 \, d\underline{x}$	$\simeq h_T^{-1}$	$\simeq h_T^d$	1	$\simeq h_T^{d-1}$
$\displaystyle\int_{\partial T} |b_n| (u^{up})^2 \, d\underline{s}$	1	$\simeq h_T^{d-1}$	1	$\simeq h_T^{d-1}$
$\displaystyle\int_{\partial T \cap \Gamma_{out}} |b_n| u_F^2 \, d\underline{s}$	1	$\simeq h_T^{d-1}$	1	$\simeq h_T^{d-1}$

So the last step of the proof of Proposition 1.2.6 reads

$$\left(\|(\mathbf{id}_h^T - \mathbf{\Pi}_{h,k}^T)u\|_C^T \right)^2 \preceq \|b\|_\infty h_T^d\, h_T^{-1}\, h_T^{-d}\, h_T^{2m}\, |u|_{H^m(T)}^2 = \|b\|_\infty h_T^{2k+1} |u|_{H^m(T)}^2 \tag{1.3.34}$$

which proves the claim.

$\square$

1.3.2.5   Putting it all together

HDG norm estimates

Lemma 1.3.6 (Modified Cea's Lemma). *Let* $\mathbf{u}_h \in V_h^k$ *be the solution of* (1.3.5) *and* $u \in H^1(\Omega) \cap H^m(\mathcal{T}_h) \subset W$, $m \geq 2$ *the solution of (1.3.1). Then the error in the* $\|\cdot\|_{C,*}$*-norm is bounded only a constant away from the* approximation error *in the* $\|\cdot\|_C$*-norm:*

$$\|\mathbf{u} - \mathbf{u}_h\|_{C,*} \leq C \inf_{\mathbf{v}_h \in V_{h,D}} \|\mathbf{u} - \mathbf{v}_h\|_C \tag{1.3.35}$$

for $C \in \mathbb{R}$ *only depending on the shape regularity.*

Proof. First of all we can show that the difference of a function $\mathbf{v}_h \in V_h$ to the discrete solution measured in the $\|\cdot\|_{C,*}$-norm can be bounded by the difference to the exact solution in the $\|\cdot\|_C$-norm

$$\|\mathbf{v}_h - \mathbf{u}_h\|_{C,*} \leq \frac{1}{\alpha_{\mathcal{C}_h}} \|\mathbf{v}_h - \mathbf{u}\|_C \tag{1.3.36}$$

This holds with inf-sup stability, consistency and boundedness of $\mathcal{C}_h$ ($\tilde{\mathbf{v}} \in V_{h,0}$ is again as in Prop. 1.3.2):

$$\begin{aligned}
\|\mathbf{v}_h - \mathbf{u}_h\|_{C,*} \cdot \|\tilde{\mathbf{v}}\|_{C,*} &\leq \frac{1}{\alpha_{\mathcal{C}_h}} \mathcal{C}_h(\mathbf{v}_h - \mathbf{u}_h, \tilde{\mathbf{v}}) \\
&= \frac{1}{\alpha_{\mathcal{C}_h}} \Big(\mathcal{C}_h(\mathbf{v}_h - \mathbf{u}, \tilde{\mathbf{v}}) + \underbrace{\mathcal{C}_h(\mathbf{u} - \mathbf{u}_h, \tilde{\mathbf{v}})}_{=0} \Big) \\
&\leq \frac{1}{\alpha_{\mathcal{C}_h}} \|\mathbf{v}_h - \mathbf{u}\|_C \cdot \|\tilde{\mathbf{v}}\|_{C,*}
\end{aligned} \tag{1.3.37}$$

So after separating the *discrete error* and the *approximation error* and using (1.3.36) and (1.3.25) we get

$$\|\mathbf{u} - \mathbf{u}_h\|_{C,*}^2 \leq \|\mathbf{v}_h - \mathbf{u}_h\|_{C,*}^2 + \|\mathbf{v}_h - \mathbf{u}\|_{C,*}^2 \leq \Big(\frac{1}{\alpha_{\mathcal{C}_h}^2} + C_2\Big)\|\mathbf{v}_h - \mathbf{u}\|_C^2 \tag{1.3.38}$$

and so the claim holds for the constant $C = \sqrt{\frac{1}{\alpha_{\mathcal{C}_h}^2} + C_2}$ with C_2 the constant of (1.3.25). $\qquad\square$

As mentioned before, $H^m(\mathcal{T}_h)$ is normally to much to demand from a pure hyperbolic problem. However, it is the easiest way to get error estimates in the HDG norm:

Proposition 1.3.7 (HDG norm estimate). *Let* $\mathcal{T}_h$ *be a quasi-uniform shape regular mesh,* $\mathbf{u}_h \in V_h^k$ *be the solution of* (1.3.5) *and* $u \in H^1(\Omega) \cap H^m(\mathcal{T}_h) \subset W$, $m \geq 2$ *the solution of (1.3.1). Then there holds the following error estimate:*

$$\|\mathbf{u} - \mathbf{u}_h\|_{C,*} \leq C \sqrt{\|\underline{b}\|}_\infty h^s |u|_{H^m(\mathcal{T}_h)} \qquad s = \min(k + \frac{1}{2}, m - \frac{1}{2}) \tag{1.3.39}$$

with $C \in \mathbb{R}$ *only depending on the shape regularity.*

Proof. The claim follows with Proposition 1.3.5 and Lemma 1.3.6. $\qquad\square$

Remark 1.3.2 (L^2 norm estimates):
Although in principal some kind of *Aubin Nitsche* technique is possible to gain convergence results in the L^2 norm, the requirements on the adjoint problem are far from realistic, at least when working with $H^m(\mathcal{T}_h)$. So instead of a proof of convergence in the L^2 norm, it should just be mentioned that under "good conditions" the L^2-norm is expected to convergence one half order faster than the $\|\cdot\|_{C,*}$-norm. However, there exist examples where the convergence rate $h^{k+\frac{1}{2}}$ in the L^2 norm is the best you can get even for smooth solutions.

1.3.3   Numerical Example: A steady rotating convection problem

This example is taken and modified from [NPC09]. We consider the domain $\Omega = (0,1)^2 \setminus \Gamma$ with $\Gamma = \frac{1}{2} \times (0, \frac{1}{2})$ and the convective velocity $\underline{b} = \begin{pmatrix} y - \frac{1}{2} \\ \frac{1}{2} - x \end{pmatrix}$ which is divergence-free. On all inflow boundaries of $(0,1)^2$ Dirichlet boundary conditions $u = 0$ and $u = 1 - \tanh(10(1 - 4y))$ on the left (inflow) part of Γ are prescribed.

(a) Setup of the problem: red boundaries are inflow boundaries, blue boundaries are outflow boundaries

(b) Initial mesh (Level 0) and discrete solution ($p = 5$)
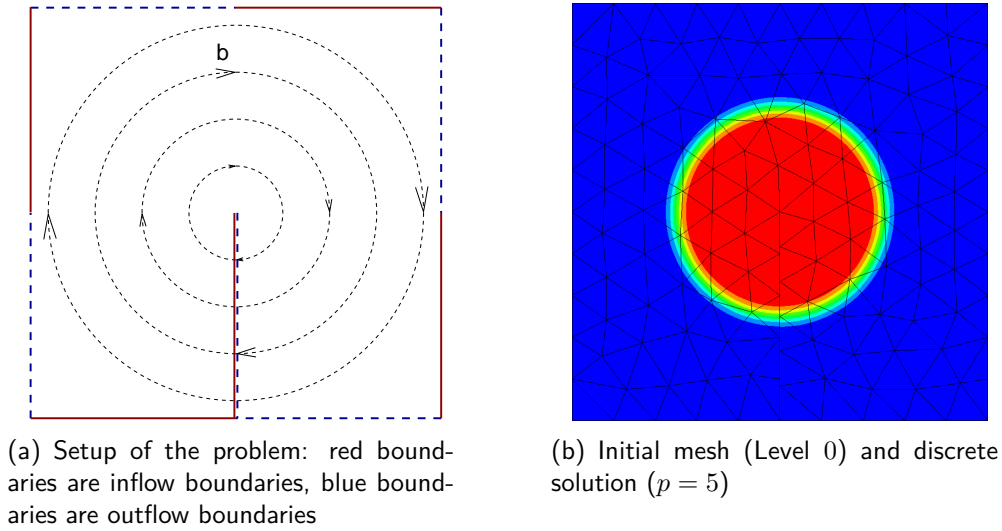
Figure 1.3.3: Setup, initial mesh and solution of Numerical Example 1.3.3

The Dirichlet boundary condition should be transported from the left to the right side of Γ.
Due to non-aligned meshes the discrete solution will be diffusive on coarse meshes, i.e. the profile on the right side of the slit will be smeared out a little bit. For different spatial and polynomial resolutions we plotted the exact solution, the prescribed profile left of the slit and the discrete solution of the profile right of the slit in Figure 1.3.4. As our solution is smooth, higher polynomial degree reduces this smearing effect faster as higher grid resolution. We also see in the following table that optimal convergence rates in the L^2-norm are achieved, what indicates that for smooth functions the results of our analysis are often pessimistic. Notice that as gradients of the solution are steep, high resolution is necessary before we observe the asymptotically correct convergence rate. See e.g. that for $p = 0$ the resolution after six uniform refinement steps is still too poor to get order 1 convergence.

meshlevel (elements)	p	unknowns DOF	$\|\mathbf{u} - \mathbf{u}_h\|_{L^2}$ error	order	p	unknowns DOF	$\|\mathbf{u} - \mathbf{u}_h\|_{L^2}$ error	order
0 (231)	0	603	0.3813	—	1	1 437	0.1231	—
1 (924)		2 361	0.2648	0.526		5 646	0.07563	0.703
2 (3 696)		9 342	0.1956	0.437		22 380	0.02667	1.504
3 (14 784)		37 164	0.1352	0.533		89 112	7.560e-3	1.819
4 (59 136)		148 248	0.0894	0.598		355 632	1.539e-3	2.296
5 (236 544)		592 176	0.0558	0.678		1 420 896	2.861e-4	2.428
6 (946 176)		2 367 072	0.0329	0.764				
0 (231)	3	3 798	0.0252	—	5	7 083	4.610e-3	—
1 (924)		14 988	8.517e-3	1.568		28 026	8.396e-4	2.457
2 (3 696)		59 544	6.981e-4	3.609		111 492	3.747e-5	4.486
3 (14 784)		237 360	4.015e-5	4.120		444 744	5.985e-7	5.968
4 (59 136)		947 808	2.483e-6	4.015		1 776 528	8.484e-9	6.141
5 (236 544)		3 787 968	1.558e-7	3.994				

Table 1.2: L^2-Error-Convergence of the HDG method for the rotating convection problem

(a) $p = 0$, mesh level 4

(b) $p = 1$, mesh level 2

(c) $p = 3$, mesh level 0

(d) $p = 5$, mesh level 0
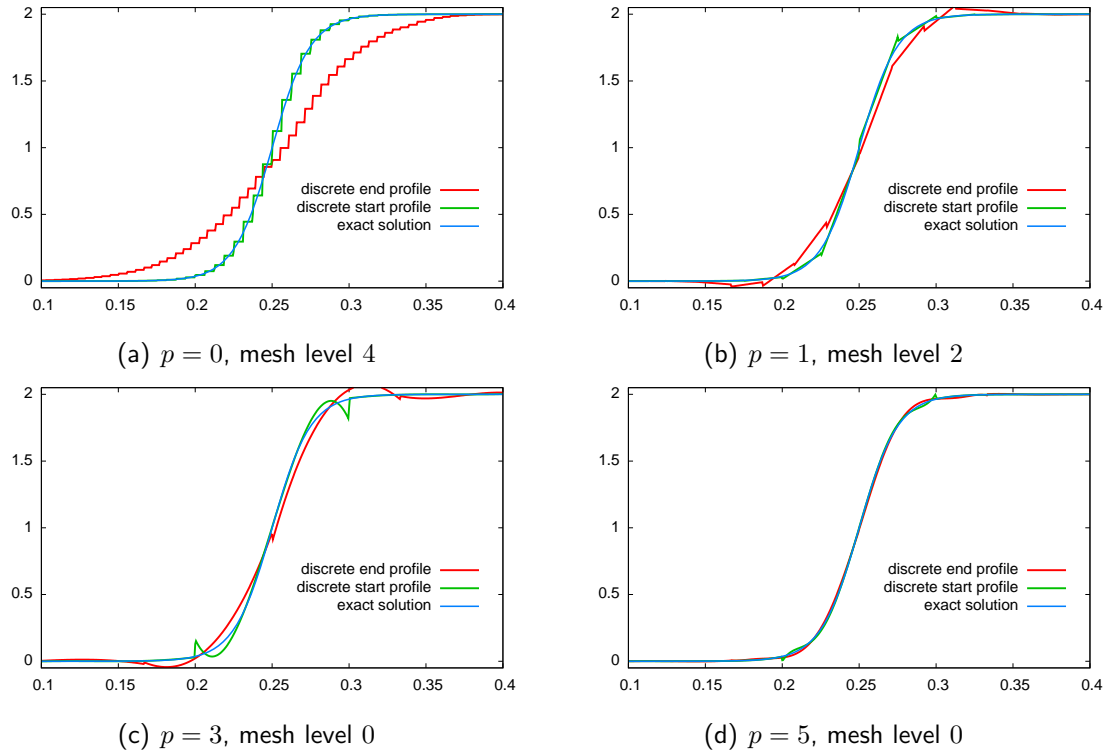
Figure 1.3.4: Both sides of the slit for different polynomial and grid resolutions compared to exact solution

1.4   Adding diffusion and convection together

Subject of this section is now the steady version of the original model problem (1.1.1)

$$
\begin{cases}
\operatorname{div}(-\varepsilon\underline{\nabla}u + \underline{b}u) & = f & \text{in } \Omega, \quad \operatorname{div}(\underline{b}) = 0 \\
u & = u_D & \text{on } \Gamma_D \\
\frac{\partial u}{\partial \underline{n}} & = g & \text{on } \Gamma_N \\
\frac{\partial u}{\partial \underline{n}} + \beta u & = h & \text{on } \Gamma_R
\end{cases}
\tag{1.4.1}
$$

with $\varepsilon > 0$ and $\|\underline{b}\| \geq 0$ everywhere.

1.4.1   Joining the limits

The discretization of the convection diffusion problem is easy as the approximation spaces of both subproblems are equal. The appropriate generalization of the proposed formulations is achieved by adding both bilinearforms together:

$$
\mathcal{A}_h(\mathbf{u}, \mathbf{v}) := \mathcal{B}_h(\mathbf{u}, \mathbf{v}) + \mathcal{C}_h(\mathbf{u}, \mathbf{v})
\tag{1.4.2}
$$

and the discrete problem now reads:

$$
\begin{aligned}
&\text{Find } \mathbf{u} \in V_D, \text{ such that for all } \mathbf{v} \in V_{h,0} : \\
&\mathcal{A}_h(\mathbf{u}, \mathbf{v}) = \langle f, v \rangle
\end{aligned}
\tag{1.4.3}
$$

Remark 1.4.1 (Comparison to Standard DG):
Notice that although in both limit cases the bilinearforms could be transformed to Standard DG bilinearform $\mathcal{B}_h^{DG}$ and $\mathcal{C}_h^{DG}$ respectively, the corresponding elimination of the facet unknowns in the general case does not coincide with the sum of both Standard DG bilinearforms.

$$
\mathcal{A}_h^{DG}(\mathbf{u}, \mathbf{v}) \neq \mathcal{B}_h^{DG}(\mathbf{u}, \mathbf{v}) + \mathcal{C}_h^{DG}(\mathbf{u}, \mathbf{v})
\tag{1.4.4}
$$

This becomes immediately clear when you look at the result of the elimination of u_F in the next section.

1.4.1.1   Characterization of the numerical trace

We motivated and derived convenient choices for the numerical flux in both limit cases $\|\underline{b}\| \to 0$ and $\varepsilon \to 0$. In the elliptic case we have seen that we use a central approximation which is stabilized by some stabilization term depending on the discontinuity of the flux, whereas in the pure hyperbolic case we used the upwind choice which is known to be the natural and stable choice. When both, diffusion and convection are considered, the numerical trace lies in between those choices. Here we want to briefly describe how the intermediate numerical trace looks like. If we again only test with facet functions v_F we can reconstruct an expression for the facet value u_F:

$$
u_F = \frac{2\{\!\!\{\tau_h\varepsilon u\}\!\!\} - [\![\varepsilon\frac{\partial u}{\partial \underline{n}}]\!] + |b_n|u^{DG}}{2\{\!\!\{\tau_h\varepsilon\}\!\!\} + |b_n|}
\tag{1.4.5}
$$

Thus the conserved total flux J_h is describes as

$$
J_h := b_n u^{up} - \varepsilon\frac{\partial u}{\partial \underline{n}} + \varepsilon\tau_h(u - u_F)
\tag{1.4.6}
$$

1.4.2 A priori error analysis

When adding diffusion to the convection, problems with the regularity vanish (at least from a purist point of view). The solution of the convection diffusion equation has the same regularity as the same problem without convection. Nevertheless the effect of "nearly discontinuities" may appear (e.g. at boundaries) which renders the apperent regularity useless in many cases.

Consistency and boundedness in the a priori error analysis follow straight forward from the limiting problems. For stability we have to put some extra technical effort into the proof. The approximation section is in general easily reconstructed from the elliptic case, but some extra attention has to be taken as the regularity may degenerate.

Again we assume to have Dirichlet data on the whole boundary.

1.4.2.1 HDG norms

The appropriate *HDG norms* for the convection diffusion problem also follow directly from the appropriate norms of the subproblems.

$$\|\mathbf{u}\|_*^2 = \|\mathbf{u}\|_{C,*}^2 + \|\mathbf{u}\|_{\varepsilon,*}^2 \qquad \|\mathbf{u}\|^2 = \|\mathbf{u}\|_C^2 + \|\mathbf{u}\|_\varepsilon^2 \tag{1.4.7}$$

In both limits the *HDG norms* coincide with the appropriate of the limit problem.

1.4.2.2 Consistency

Consistency follows directly from the consistency of the pure hyperbolic and the pure elliptic problem. Let $u \in V_D$ be the solution of (1.4.1), then for all $v \in V_0$

$$
\begin{aligned}
\mathcal{A}_h(\mathbf{u}, \mathbf{v}) &= \sum_{T \in \mathcal{T}_h} \left\{ \int_T \mathrm{div}(\underline{b}u)v \, d\underline{x} + \underbrace{\int_{\partial T_{in}} |b_n|(u - u_F) \ v \, d\underline{s}}_{=0} + \underbrace{\int_{\partial T_{out}} |b_n|(u_F - u) \ v_F \, d\underline{s}}_{=0} \right. \\
&\qquad \left. + \int_T \varepsilon \, \underline{\nabla}u \, \underline{\nabla}v \, d\underline{x} - \int_{\partial T} \varepsilon \, \frac{\partial u}{\partial \underline{n}} \, (v - v_F) \, d\underline{s} \right\} \\
&= \sum_{T \in \mathcal{T}_h} \left\{ \int_T \underbrace{\mathrm{div}(\underline{b}u - \varepsilon \underline{\nabla}u)}_{=f} v \, d\underline{x} + \int_{\partial T} \varepsilon \, \frac{\partial u}{\partial \underline{n}} \, v_F \, d\underline{s} \right\} \\
&= \int_\Omega f \ v \, d\underline{x} + \sum_{E \in \mathcal{F}_h^{int}} \int_E \underbrace{[\![\varepsilon \, \underline{\nabla}u]\!]}_{=0} \, v_F \, d\underline{s} + \sum_{E \in \mathcal{F}_h^{ext}} \int_E \varepsilon \frac{\partial u}{\partial n} \, \underbrace{v_F}_{=0} \, d\underline{s}
\end{aligned}
\tag{1.4.8}
$$

Proposition 1.4.1 (Galerkin Orthogonality). *Let $\mathbf{u}_h \in V_{h,D}$ be the solution of* (1.4.3)

$$\mathcal{A}_h(\mathbf{u}_h, \mathbf{v}_h) = (f, v_h)_\Omega \qquad \forall \ \mathbf{v}_h \in V_{h,0}$$

and u be the solution of (1.4.1). *Then there holds*

$$\mathcal{A}_h(\mathbf{u}_h - \mathbf{u}, \mathbf{v}_h) = 0 \qquad \forall \ \mathbf{v}_h \in V_{h,0} \tag{1.4.9}$$

Proof. Using (1.4.8) we have: For all $\mathbf{v}_h \in V_h$

$$\mathcal{A}_h(\mathbf{u} - \mathbf{u}_h, \mathbf{v}_h) = \int_\Omega f \ v_h \, d\underline{x} - \int_\Omega f \ v_h \, d\underline{x} = 0 \tag{1.4.10}$$

$\square$

1.4.2.3 Stability

Proposition 1.4.2 (inf-sup condition). *Let u be the exact solution of (1.4.1). Then, for a constant $\alpha_{\mathcal{A}_h} \in \mathbb{R}$ independent of the meshsize h and $\tau_h h$ sufficiently large the following holds:*

$$\sup_{\mathbf{v} \in V_{h,0}} \frac{\mathcal{A}_h(\mathbf{u}, \mathbf{v})}{\|\mathbf{v}\|_*} \geq \alpha_{\mathcal{A}_h} \|\mathbf{u}\|_* \quad \forall\, \mathbf{u} \in V_{h,0} \tag{1.4.11}$$

Furthermore for every $\mathbf{u} \in V_{h,0}$ we can find a $\tilde{\mathbf{v}} \in V_{h,0}$, s.t.

$$\mathcal{A}_h(\mathbf{u}, \tilde{\mathbf{v}}) \geq \alpha_{\mathcal{A}_h} \|\mathbf{u}\|_* \|\tilde{\mathbf{v}}\|_* \tag{1.4.12}$$

Proof. Similar to the approach we used in pure convective case when showing inf-sup stability in section 1.3.2.2 we will test $\mathcal{A}_h(\cdot, \cdot)$ with $\tilde{\mathbf{v}} = \gamma \mathbf{u} + \mathbf{v}^*$, with $\gamma \geq 2c_{sr}^2 + \frac{1}{2}$. We can bound the elliptic HDG norm of the last part $\mathbf{v}^* = (\frac{h}{|\underline{b}|}\underline{b} \cdot \nabla u, 0)$ also with $\|\mathbf{u}\|_{\varepsilon,*}$ when using scaling arguments $(\|\nabla w\|_T \leq \frac{\tilde{c}_k}{h^2}\|w\|_T^2 \ \forall\, w \in \mathcal{P}^k(T))$ and an inverse inequality:

$$
\begin{aligned}
\|\mathbf{v}^*\|_{\varepsilon,*}^2 \ &:= \ \sum_{T \in \mathcal{T}_h} \varepsilon \left\{ \left\| \frac{h}{|\underline{b}|}\nabla(\underline{b} \cdot \nabla u) \right\|_T^2 + \frac{1}{h}\left\| \frac{h}{|\underline{b}|}(\underline{b} \cdot \nabla u) \right\|_{\partial T}^2 \right\} \\
&\leq \ \sum_{T \in \mathcal{T}_h} \varepsilon \left\{ h^2 \|\nabla(\nabla u)\|_T^2 + h \|\nabla u\|_{\partial T}^2 \right\} \\
&\leq \ (\tilde{c} + c_{sr}^2) \sum_{T \in \mathcal{T}_h} \varepsilon \|\nabla u\|_T^2 \qquad \leq (\tilde{c} + c_{sr}^2)\|\mathbf{u}\|_{\varepsilon,*}^2
\end{aligned} \tag{1.4.13}
$$

Keeping that in mind and using (1.3.21), it follows

$$\|\tilde{\mathbf{v}}\|_* \leq k\|\mathbf{u}\|_* \tag{1.4.14}$$

where k depends linearly on c, the constant of the shape regularity, and on fixed (problem-independent) constants and γ. With the last inequality it is sufficient to show

$$\mathcal{A}_h(\mathbf{u}, \tilde{\mathbf{v}}) \geq K\|\mathbf{u}\|_*^2 \qquad \text{for some constant } K \in \mathbb{R} \tag{1.4.15}$$

The last ingredient to show this is to bound $\mathcal{B}_h(\mathbf{u}, \mathbf{v}^*)$ as the other arising terms were already bounded in the analysis of the pure hyperbolic and the pure elliptic case, respectively. The bound follows again

with Cauchy-Schwarz inequality and the compiled relations from above:

$$
\begin{aligned}
\mathcal{B}_h(\mathbf{u}, \mathbf{v}^*) \quad :=\quad & \sum_{T \in \mathcal{T}_h} \left\{ \int_T \varepsilon \, \underline{\nabla} u \, \underline{\nabla} v^* \, d\underline{x} - \int_{\partial T} \varepsilon \, \frac{\partial u}{\partial \underline{n}} \, [\![\mathbf{v}^*]\!] \, d\underline{s} \right. \\
& \left. - \int_{\partial T} \varepsilon \, \frac{\partial v^*}{\partial \underline{n}} \, [\![\mathbf{u}]\!] \, d\underline{s} + \int_{\partial T} \varepsilon \tau_h \, [\![\mathbf{u}]\!] \, [\![\mathbf{v}^*]\!] \, d\underline{s} \right\} \\
\overset{C.S.}{\geq} \quad & -\sum_{T \in \mathcal{T}_h} \varepsilon \left\{ \|\underline{\nabla} u\|_T \|\underline{\nabla} v^*\|_T + h^{\frac{1}{2}} \left\| \frac{\partial u}{\partial \underline{n}} \right\|_{\partial T} h^{-\frac{1}{2}} [\![\mathbf{v}^*]\!]_{\partial T} \right. \\
& \left. + h^{\frac{1}{2}} \left\| \frac{\partial v^*}{\partial \underline{n}} \right\|_{\partial T} h^{-\frac{1}{2}} [\![\mathbf{u}]\!]_{\partial T} + \tau_h [\![\mathbf{u}]\!]_{\partial T} [\![\mathbf{v}^*]\!]_{\partial T} \right\} \\
\overset{C.S.}{\geq} \quad & -\sum_{T \in \mathcal{T}_h} \varepsilon \left\{ \left(\|\underline{\nabla} u\|_T^2 + h \left\| \frac{\partial u}{\partial \underline{n}} \right\|_{\partial T}^2 + \frac{1}{h} [\![\mathbf{u}]\!]_{\partial T}^2 + \tau_h [\![\mathbf{u}]\!]_{\partial T}^2 \right)^{\frac{1}{2}} \right. \\
& \left. \left(\|\underline{\nabla} v^*\|_T^2 + h \left\| \frac{\partial v^*}{\partial \underline{n}} \right\|_{\partial T}^2 + \frac{1}{h} [\![\mathbf{v}^*]\!]_{\partial T}^2 + \tau_h [\![\mathbf{v}^*]\!]_{\partial T}^2 \right)^{\frac{1}{2}} \right\} \\
\overset{\text{inv. ineq.}}{\geq} \quad & -\sum_{T \in \mathcal{T}_h} \max(1 + c_T^2, 1 + \tau_h h) \|\mathbf{u}\|_{\varepsilon,*}^T \|\mathbf{v}^*\|_{\varepsilon,*}^T \\
\geq \quad & -\max\left(1 + c_{sr}^2, \sup_{\underline{x} \in \mathcal{F}_h} (1 + \tau_h h) \right) \|\mathbf{u}\|_{\varepsilon,*} \|\mathbf{v}^*\|_{\varepsilon,*} \\
\overset{(1.4.14)}{\geq} \quad & -K_2 \|\mathbf{u}\|_{\varepsilon,*}^2 \qquad \text{with } K_2 = \max\left(1 + c_{sr}^2, \sup_{\underline{x} \in \mathcal{F}_h} (1 + \tau_h h) \right) \cdot k
\end{aligned}
$$
(1.4.16)

Finally we are able to show the inf-sup condition. We choose the proposed test function but additionally require that $\gamma \geq \frac{K_2 + \frac{1}{2}}{\alpha_{\mathcal{B}_h}}$. Thereby we achieve the following:

$$
\mathcal{A}_h(\mathbf{u}, \tilde{\mathbf{v}}) = \gamma \underbrace{\mathcal{B}_h(\mathbf{u}, \mathbf{u})}_{\geq \alpha_{\mathcal{B}_h} \|\mathbf{u}\|_{\varepsilon,*}^2} + \underbrace{\mathcal{B}_h(\mathbf{u}, \mathbf{v}^*)}_{\geq -K_2 \|\mathbf{u}\|_{\varepsilon,*}^2} + \underbrace{\mathcal{C}_h(\mathbf{u}, \tilde{\mathbf{v}})}_{\geq \frac{1}{2} \|\mathbf{u}\|_{C,*}^2} \geq \frac{1}{2} \|\mathbf{u}\|_*^2 \overset{(1.4.14)}{\geq} \underbrace{\frac{1}{2k}}_{\alpha_{\mathcal{A}_h}} \|\mathbf{u}\|_* \|\tilde{\mathbf{v}}\|_*
$$
(1.4.17)

where we used coercivity of $\mathcal{B}_h$ and equations (1.3.20) and (1.4.16). $\qquad \square$

1.4.2.4 Boundedness

Boundedness (Continuity) follows directly from the continuity of the subproblems

Proposition 1.4.3. *For all $\mathbf{u}, \mathbf{v} \in V$ there holds:*

$$
|\mathcal{A}_h(\mathbf{u}, \mathbf{v})| \leq \beta_{\mathcal{A}_h} \|\mathbf{u}\|_C \|\mathbf{v}\|_{C,*}
$$

with $\beta_{\mathcal{A}_h} = \beta_{\mathcal{B}_h} = \sup_{\underline{x} \in \mathcal{F}_h} (1 + \tau_h h)$

Proof. We just have to use the boundedness results for $\mathcal{B}_h$ and $\mathcal{C}_h$ and directly get

$$
\begin{aligned}
\mathcal{A}_h(\mathbf{u}, \mathbf{v}) \quad = \quad & \mathcal{B}_h(\mathbf{u}, \mathbf{v}) + \mathcal{C}_h(\mathbf{u}, \mathbf{v}) \\
\leq \quad & \|\mathbf{u}\|_{\varepsilon,*} \|\mathbf{v}\|_\varepsilon + \sup_{\underline{x} \in \mathcal{F}_h} (1 + \tau_h h) \|\mathbf{u}\|_{C,*} \|\mathbf{v}\|_C \\
\overset{C.S.}{\leq} \quad & \sup_{\underline{x} \in \mathcal{F}_h} (1 + \tau_h h) \|\mathbf{u}\|_* \|\mathbf{v}\|
\end{aligned}
$$
(1.4.18)

$\qquad \square$

1.4.2.5   Approximation

Plugging in the results of the pure convective and the pure elliptic case directly gives

Proposition 1.4.4. *On a shape regular mesh $\mathcal{T}_h$ which consists of affine transformed elements and a function $u \in H^1(\Omega) \cap H^m(\mathcal{T}_h)$, $m = k + 1 \geq 2$ there holds the following approximation result for V_h^k as introduced in* (1.2.35c)

$$\inf_{\mathbf{v}_h \in V_h^k} \|\mathbf{u} - \mathbf{v}_h\|^2 \leq \sum_{T \in \mathcal{T}_h} \left(C_1 \|\underline{b}\|_\infty h_T^{2k+1} + C_2 \|\varepsilon\|_\infty h_T^{2k} \right) |u|_{H^m(T)}^2 \tag{1.4.19}$$

with $C1$, $C2$ independent of the mesh size. For quasi-uniform meshes the direct conclusion of it is

$$\inf_{\mathbf{v}_h \in V_h^k} \|\mathbf{u} - \mathbf{v}_h\|^2 \leq \left(C_1 \|\underline{b}\|_\infty h^{2k+1} + C_2 \|\varepsilon\|_\infty h^{2k} \right) |u|_{H^m(\mathcal{T}_h)}^2 \tag{1.4.20}$$

Proof. The claim follows with the definition of the *HDG norm* and the approximation results for the limit cases of the convection diffusion equation. □

1.4.2.6   Putting it all together

HDG norm estimates

Lemma 1.4.5. *Let $\mathbf{u}_h \in V_h^k$ be the solution of* (1.4.3) *and $u \in H^1(\Omega) \cap H^m(\mathcal{T}_h) \subset W$, $m \geq 2$ the solution of* (1.3.1). *Then the error in the $\|\cdot\|_*$-norm is only bounded by a constant and the approximation error in the $\|\cdot\|$-norm:*

$$\|\mathbf{u} - \mathbf{u}_h\|_* \leq C \inf_{\mathbf{v}_h \in V_{h,D}} \|\mathbf{u} - \mathbf{v}_h\| \tag{1.4.21}$$

for $C \in \mathbb{R}$ only depending on the shape regularity and the stabilization parameter τ_h.

Proof. The proof works exactly the same way as the one of Lemma 1.3.6, just exchange $\mathcal{C}_h$ by $\mathcal{A}_h$ and adapt the constants. □

Now it is reasonable to assume $H^m(\mathcal{T}_h)$, at least for a small mesh Peclet number $Pe_h = \frac{|b|h}{\varepsilon}$ and we can work with the following error estimate:

Proposition 1.4.6 (HDG norm estimate). *Let $\mathcal{T}_h$ be a quasi-uniform shape regular mesh, $\mathbf{u}_h \in V_h^k$ be the solution of* (1.4.3) *and $u \in H^1(\Omega) \cap H^m(\mathcal{T}_h) \subset W$, $m \geq 2$ the solution of* (1.4.1). *Then there holds the following error estimate:*

$$\|\mathbf{u} - \mathbf{u}_h\|_{C,*} \leq (C_1 \sqrt{\|\underline{b}\|}_\infty \sqrt{h} + C_2 \sqrt{\|\varepsilon\|}_\infty) \, h^s |u|_{H^m(\mathcal{T}_h)} \qquad s = \min(k, m-1) \tag{1.4.22}$$

with $C \in \mathbb{R}$ only depending on the shape regularity.

Proof. The claim follows with Proposition 1.4.4 and Lemma 1.4.5. □

Remark 1.4.2 (L^2 norm estimates):
In general L^2 norm estimates are as difficult to get as in the pure convective limit. But if we assume small Peclét numbers, that is convection is not dominant anywhere, we can also get optimal order L^2 norm estimates. The easiest way to see that, is to remember (1.3.18) and thus the convection bilinearform $\mathcal{C}_h$ does not destroy the coercivity of $\mathcal{A}_h$ in the $\|\cdot\|_\varepsilon$ norm. Additionally we have to add a boundedness estimate for $\mathcal{A}_h$ which uses $\|\cdot\|_\varepsilon$ norms only. This estimate has order one constants as long as the Peclét numbers are small. In this framework we could easily adapt the results of the pure elliptic case to come to similar results and thereby to optimal order L^2 norm a priori estimates also for the convection diffusion equation. Nevertheless, the convection dominated case is often the more interesting and more challenging case. The analysis carried out in this section is still valid in the convection dominated case. When we turn over to *Navier-Stokes equations* or more specifically to the *Oseen problem*, things get more complicated and we will use the kind of analysis which makes use of the coercivity, instead of the one we applied in this chapter.

Remark 1.4.3 (A posteriori error estimates):
Although a posteriori error estimators are not being discussed here, due to the similarity to Standard DG methods, adapted versions of the error estimators proposed in [SZ09] would be possible.

1.4.3 Numerical Example: A steady convection-dominated problem

We take the numerical test of [ES09, Example 1] where the square $(0,1)^2$ with homogeneous dirichlet boundary conditions is considered. The flow field is constant $\underline{b} = (b_1, b_2)^T$, the diffusion coefficient ε is constant and the source term f is set to be:

$$f = b_1 \left(y - \frac{e^{b_2 y/\varepsilon} - 1}{e^{b_2/\varepsilon} - 1} \right) + b_2 \left(x - \frac{e^{b_1 y/\varepsilon} - 1}{e^{b_1/\varepsilon} - 1} \right) \tag{1.4.23}$$

Then, for $\varepsilon > 0$, the exact solution to the convection-diffusion problem (1.4.1) is

$$u(x, y) = \left(x - \frac{e^{b_1 x/\varepsilon} - 1}{e^{b_1/\varepsilon} - 1} \right) \left(y - \frac{e^{b_2 y/\varepsilon} - 1}{e^{b_2/\varepsilon} - 1} \right) \tag{1.4.24}$$

For our numerical tests we take $\varepsilon = 0.01$ and $\underline{b} = (2, 1)^T$. The solution has boundary layers at the outflow boundaries, i.e. the top and right side (cp. Figure 1.4.1(a)). We want to compare the proposed method with the results given in [ES09, Example 1], i.e. with the streamline diffusion method and the mixed Hybrid DG method of [ES09]. For our Hybrid DG method we again choose

$$\tau_h = \alpha \frac{(p+1)(p+d)}{d} \frac{|\partial T|}{|T|}$$

with $\alpha = 2$.

Notice that $p = 0$ is not possible for our proposed method. For $p = 1$ and $p = 2$ the results for our method are quite similar to the ones of the mixed HDG method. For all method we can not expect optimal convergence rates on the first refinement steps as the boundary layers are not yet resolved. Nevertheless, we can observe some important properties: In contrast to both HDG methods, the streamline diffusion method introduces large layers which stem from the stabilization of the convective term, so most of the error is owed to the overdiffusive stabilization. Because of that the gap between discrete solutions and best approximation(see Table 1.4 left) increases for higher k, whereas both HDG methods are very close to their L^2 best approximation(see Table 1.4 right).

---

[14] $\tilde{V}_h^p$ is the H^1-conforming finite element with piecewise complete polynomials of degree p

(a) Solution of the convection diffusion problem (1.4.1) for $\varepsilon = 0.01$ and $\underline{b} = (2,1)^T$

(b) Initial mesh (Level $L = 0$)
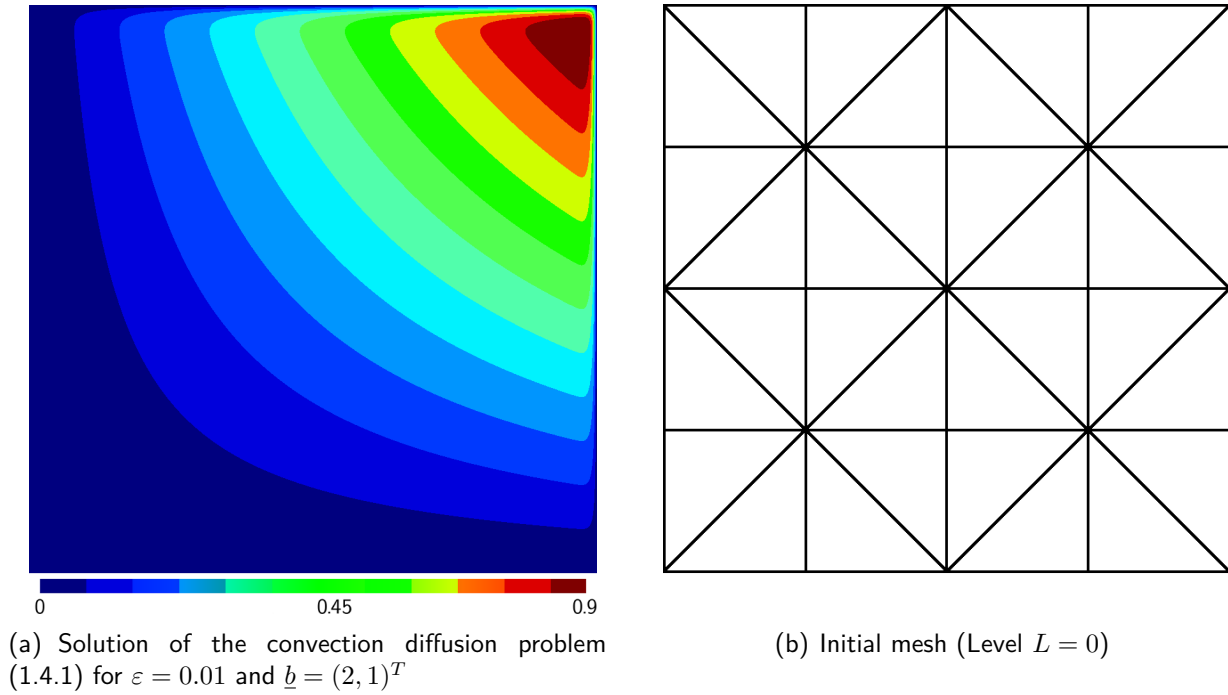
Figure 1.4.1: Initial mesh($L = 0$) and solution of Numerical Example 1.4.3

L	h	(elements)	streamline diffusion			mixed HDG ([ES09])			HDG ($\alpha = 2$)	
			$p=1$	$p=2$	$p=3$	$p=0$	$p=1$	$p=2$	$p=1$	$p=2$
0	1/4	(32)	0.21	0.19	0.19	0.065	0.040	0.033	0.040	0.034
1	1/8	(128)	0.15	0.13	0.13	0.048	0.036	0.025	0.035	0.025
2	1/16	(512)	0.097	0.088	0.088	0.040	0.026	0.014	0.025	0.014
3	1/32	(2 048)	0.056	0.050	0.050	0.032	0.014	0.0052	0.014	0.0054

Table 1.3: L^2 errors of u

1.5   Computational aspects

There are several advantages which arise with the additional facet degrees of freedom. One is the elimination of inner degrees of freedom which is possible as long as the local problems are uniquely solvable. We will show that this is the case in section 1.5.1. In section 1.5.2 the sparsity patterns of DG and HDG methods will be discussed.

L	h	(elements)	conforming[14] $\min_{v_h \in \tilde{V}_h^p} \|u - v_h\|_{L^2}$			discontinuous $\min_{v_h \in Q_h^p} \|u - v_h\|_{L^2}$		
			$p=1$	$p=2$	$p=3$	$p=0$	$p=1$	$p=2$
0	1/4	(32)	0.18	0.10	0.063	0.057	0.040	0.033
1	1/8	(128)	0.12	0.060	0.032	0.045	0.034	0.024
2	1/16	(512)	0.073	0.029	0.012	0.038	0.024	0.013
3	1/32	(2 048)	0.038	0.011	0.0030	0.030	0.013	0.0051

Table 1.4: L^2 best approximation error

1.5.1 Elimination of inner degrees of freedom

Now we want to briefly comment on the elimination of the element degrees of freedom. If we define

$$
\begin{aligned}
A(\,u\,,\,v\,) \; &= \; \sum_{T \in \mathcal{T}_h} \left\{ \int_T \varepsilon \underline{\nabla} u \underline{\nabla} v \, d\underline{x} + \int_T \mathrm{div}(\underline{b}u) v \, d\underline{x} \right. \\
&\quad \left. - \int_{\partial T} \varepsilon \left(\frac{\partial u}{\partial \underline{n}} v + \frac{\partial v}{\partial \underline{n}} u - \tau_h u v \right) d\underline{s} + \int_{\partial T_{out}} b_n u v \, d\underline{s} \right\} \\
B(\,u_F,\,v\,) \; &= \; \sum_{T \in \mathcal{T}_h} \left\{ \int_{\partial T} \varepsilon \left(\frac{\partial v}{\partial \underline{n}} u_F - \tau_h u_F v \right) d\underline{s} + \int_{\partial T_{in}} b_n u_F v \, d\underline{s} \right\} \\
C(\,u\,,\,v_F) \; &= \; \sum_{T \in \mathcal{T}_h} \left\{ \int_{\partial T} \varepsilon \left(\frac{\partial u}{\partial \underline{n}} v_F - \tau_h u v_F \right) d\underline{s} - \int_{\partial T_{out}} b_n u v_F \, d\underline{s} \right\} \\
D(\,u_F,\,v_F) \; &= \; \sum_{T \in \mathcal{T}_h} \left\{ \int_{\partial T} \varepsilon \tau_h u_F v_F \, d\underline{s} + \int_{\partial T_{out}} b_n u_F v_F \, d\underline{s} \right\}
\end{aligned}
\tag{1.5.1}
$$

we end up with the following system for u and u_F

$$
\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} u \\ u_F \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}
\tag{1.5.2}
$$

Notice that A is block diagonal and so you can invert A in an element by element fashion. Building up the schur complement (which in the elliptic case is s.p.d.[15] as $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is s.p.d.) is easy:

$$
u = A^{-1}(f - Bu_F) \Rightarrow (D - CA^{-1}B)u_F = g - CA^{-1}f
\tag{1.5.3}
$$

The solution u_F can then be used to solve for u with $Au = f - Bu_F$ element by element.

Remark 1.5.1 (Existence of local solvers):
To calculate A^{-1} we need the local element-wise problems to be uniquely solvable, that is, we have to show

$$
A(u, u) = 0 + \text{ homogeneous b.c. (in a dg sense) } \Rightarrow u \equiv 0
$$

We can do that by showing coercivity of the bilinearform A. Therefore techniques similar to those we already used to show inf-sup stability of the global bilinearform can be used. Actually solving $Au = \tilde{f}(u_F)$ for a given u_F is like solving the PDE on the small domain T where dirichlet boundary conditions are incorporated in a way, which is known as the Nitsche approach, i.e. in a weak sense.

---

[15]symmetric positive definite

Remark 1.5.2 (Alternative HDG formulation):
This property does not hold for the alternative HDG formulation presented in (see section 1.2.2.3). There we'd have to additionally prescribe some additional moment (e.g. average value) for each element to get locally solvable subproblems. Then again we would have to add those moments to the global unknowns.

1.5.2   Sparsity Pattern of HDG compared to DG methods

As we have already seen in the numerical example for HDG in the elliptic case in section 1.2.5, the efficiency of the method depends on the criteria we evaluate it with. On the one hand HDG methods, in comparison to other DG methods, come with additional degrees of freedoms, but on the other hand the stencil of the formulation is reduced and so the coupling of degrees of freedom is significantly reduced. If we also take a look at the sparsity pattern for example 1.4.3 in Figure 1.5.1(a) we observe what should already be clear from the construction of the method: element DOF don't couple directly with each other and facet DOF don't couple direclty with other facet DOF, s.t. the matrix blocks A and D of the last section are block diagonal. We can then invert each block, which is possible for A as well as for D and build according schur complements. The natural approach would be to first eliminate the element DOF (compare section 1.5.1) resulting in a linear system of equations for the facet DOF only. The corresponding sparsity pattern for the schur complement of the matrix is shown in Figure 1.5.1(a) is shown in Figure 1.5.1(b). We observe that the size of the schur complement is significantly smaller. The difference between the size of the large matrix and the size of the schur complement increases for higher polynomial degree.

(a) HDG Sparsity Pattern, Matrix has the size 2728×2728

(b) condensated HDG Sparsity Pattern, Matrix has the size 616×616

Figure 1.5.1: Sparsity Pattern of HDG and DG for the problem 1.4.3 and the initial mesh (see also Figure 1.4.1(b)) for polynomial degree $p = 10$. Dirichlet degrees of freedom are also included here.

Another possibility is to eliminate the facet DOF which results in a schur complement which has exactly the same sparsity pattern as standard dg method. This is not surprising as we already saw that

a Standard DG formulation of the hdg method can be achieved by eliminating the facet functions. The corresponding sparsity pattern can be seen in Figure 1.5.2.

We conclude that, concerning the facet functions, direct solvers as well as iterative solvers benefit more from them than they suffer from a few more degrees of freedom.

As long as we have to solve linear systems, the hybridized version of a DG methods should be prefered. Nevertheless for time-dependent problems which are solved explicitly, s.t. no linear system has to be solved, the additional degrees of freedoms on facet are an unneccessary burden.
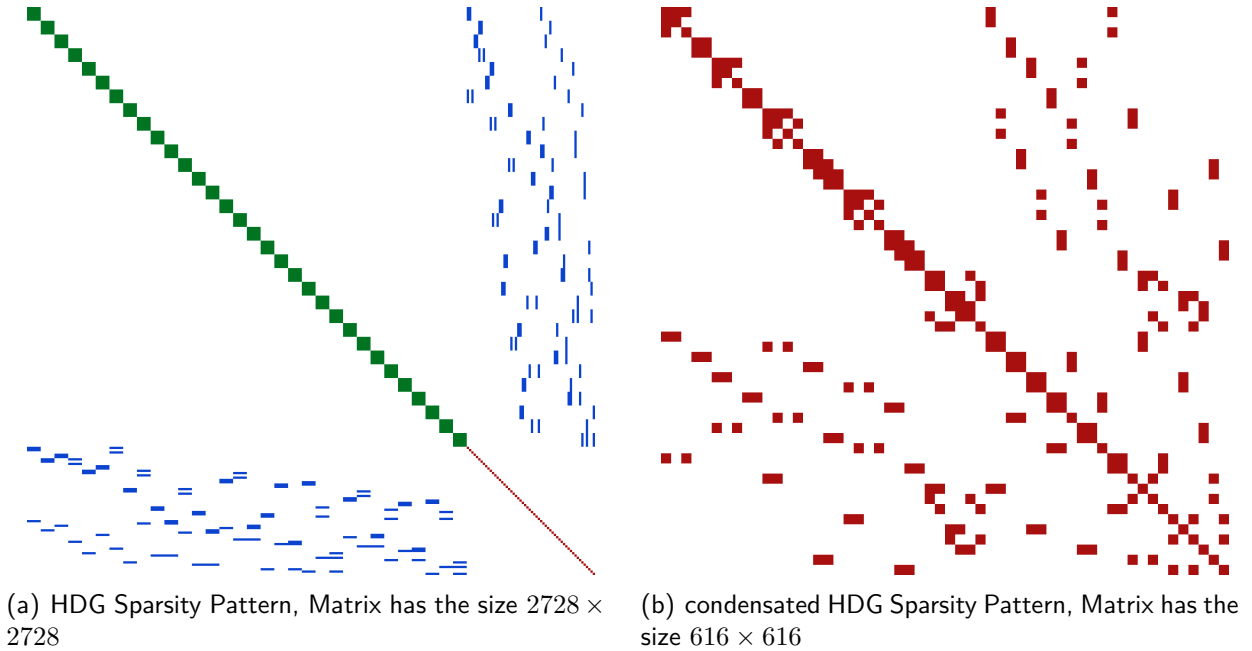

Figure 1.5.2: Sparsity Pattern of DG for the problem 1.4.3 and the initial mesh (see also Figure 1.4.1(b)) for polynomial degree $p = 10$ (Matrix has the size 2112×2112).

Chapter 2

Navier-Stokes equations with an exactly divergence-free Hybrid DG method

The (incompressible) Navier-Stokes equations describe the dynamics of an incompressible flow. They consist of one vector-valued equation for the transport of momentum and a mass conservation equation which comes as a constraint to the momentum equation. We consider two types of boundary conditions. Either the velocity is prescribed at a boundary (e.g. rigid walls or walls moving with a known velocity or inflow boundaries) which is a Dirichlet boundary condition or the momentum transport over a boundary (e.g. outflow bondaries) is known which is a Neumann boundary.

According to Newton's second law the change of momentum on a control volume balances with the sum of all forces acting on it and the in- and outflow of momentum. The outer forces sum up in $\underline{f}$, whereas the in- and outflow of momentum is decomposed into three parts:

- convective transport due to the velocity $[\quad \rho \underline{u} \otimes \underline{u} \quad]$

- viscous forces $[\quad -\eta \underline{\underline{\varepsilon}}(\underline{u}) \quad]$

- pressure forces $[\quad p \quad]$

Notice that we just consider newtonian fluids for the viscous force model, where $\underline{\underline{\varepsilon}}(\underline{u}) := \frac{1}{2}(\nabla \underline{u} + \nabla^T \underline{u})$. Often the Navier-Stokes equations are rewritten in a form which replaces $\underline{\underline{\varepsilon}}(\underline{u})$ simply by $\nabla \underline{u}$. The only difference of both formulations is actually the natural boundary conditions. The natural boundary condition for $\underline{\underline{\varepsilon}}(\underline{u}) = \nabla \underline{u}$ seems to be the more appropriate in most cases.

As we consider only incompressible flows, there holds $\rho \equiv const$. So we can devide by ρ and introduce the kinematic viscosity $\nu = \frac{\eta}{\rho}$. The volume-specific pressure p and forces $\underline{f}$ will be reinterpreted as mass-specific pressure and volume forces without change of notation from now on. Then we end up with the incompressible Navier-Stokes equations

$$\begin{cases} \frac{\partial u}{\partial t} + \text{div}(-\nu \nabla \underline{u} + \underline{u} \otimes \underline{u} + p\underline{\underline{I}}) &= \underline{f} & \text{in } \Omega \\ \text{div}(\underline{u}) &= 0 & \text{in } \Omega \\ \underline{u} &= \underline{u}_D & \text{on } \Gamma_D \\ (\nu \nabla \underline{u} - p\underline{\underline{I}}) \cdot \underline{n} &= 0 & \text{on } \Gamma_{out} \end{cases} \tag{2.0.1}$$

Our objective is to solve this set of equations numerically. There are four properties of this equations which make the numerical treatment for Navier-Stokes equations challenging: This set of equations is

1. a saddle point problem

2. vector-valued

3. nonlinear

4. unsteady

The source of the saddle point structure becomes clear if we remember that the pressure can be seen as a lagrange multiplier for the incompressibility constraint. The fact that the Navier-Stokes equations are vector-valued is not a problem of it's own but allows for distinguishing between normal and tangential component when formulating a numerical discretization. The nonlinearity of the Navier-Stokes equations is what makes them interesting: The steady problem may no longer have a unique solution and the dynamic of flow problems is often dominated by this nonlinearity. To be able to concentrate on one problem after another we will neglect the time derivative in this chapter, i.e. we just consider steady state problems. Instationarity will be the subject of chapter 3. Furthermore, we will start without the nonlinear convective term and derive a numerical method as well as the analysis for the so called Stokes problem in section 2.3. Then we will introduce a given linear convective velocity, which leads to the Oseen equations in section 2.4. Finally we will replace the convective velocity with the fluid velocity and thereby recover the Navier-Stokes equations and discuss the numerical treatment of it in section 2.5.

Before we can introduce the method, we want to propose here, for the Stokes problem, we have to make some preparations. Section 2.1 and 2.2 will motivate and introduce the use of $H(\text{div})$-conforming Finite Element methods, which will be needed for the exactly divergence-free Hybrid Discontinuous Galekin method later on.

2.1 A semi-conforming Finite Element space for Navier-Stokes

Solutions of the Navier-Stokes equations are $[H^1(\Omega)]^d$-regular under reasonable assumptions on the domain and the data. For ongoing discussions we will write $[H^1(\Omega)]^d$ in a non-standard and mesh-depend way:

$$[H^1(\Omega)]^d = \underbrace{\{v \in [H^1(\mathcal{T}_h)]^d;\ [\![v \cdot \underline{n}]\!]^* = 0 \text{ on } \mathcal{F}_h\}}_{\circledast_N} \cap \underbrace{\{v \in [H^1(\mathcal{T}_h)]^d;\ [\![v \times \underline{n}]\!]^* = 0 \text{ on } \mathcal{F}_h\}}_{\circledast_T}$$

Here we divided the continuity on element interfaces of all components into tangential- and a normal-continuity.

Our goal is to derive a method which fulfills the following properties:

1. The method should be locally and globally conservative, meaning a discrete (momentum) flux should be conserved

2. The method should be stable

3. The method should possess optimal convergence properties

To account for the first point we again consider Discontinuous Galerkin methods. Following [CKS05] a DG method which is locally conservative and energy-stable has to provide $H(\text{div}, \Omega)$-conforming solutions, that is, the discrete solution $\underline{u}$ has to be in $H(\text{div}, \Omega) = \{v \in [L^2(\Omega)]^d, \text{div}(v) \in L^2(\Omega)\}$. In this work we will use Finite Element spaces which imply this $H(\text{div}, \Omega)$-conformity for all ansatz functions. Those are taken to be piecewise polynomial and normal-continuous on element interfaces and thus are a subset of $\circledast_N$.

Tangential continuity is not included in the space so we have to account for it in another way. We will do this only in a (Hybrid) DG fashion, i.e. by adapting the formulation instead of the Finite Element space. Standard DG formulations for the tangential component of the solution are of course possible here. But, as in the scalar case, all degrees of freedom of neighbouring elements would couple directly. As well as in the scalar case, in the vector-valued case this issue can easily be overcomed with a hybridized formulation, which introduces $\underline{u}_F^t$, the approximation for the tangential trace of the solution $\underline{u}$.
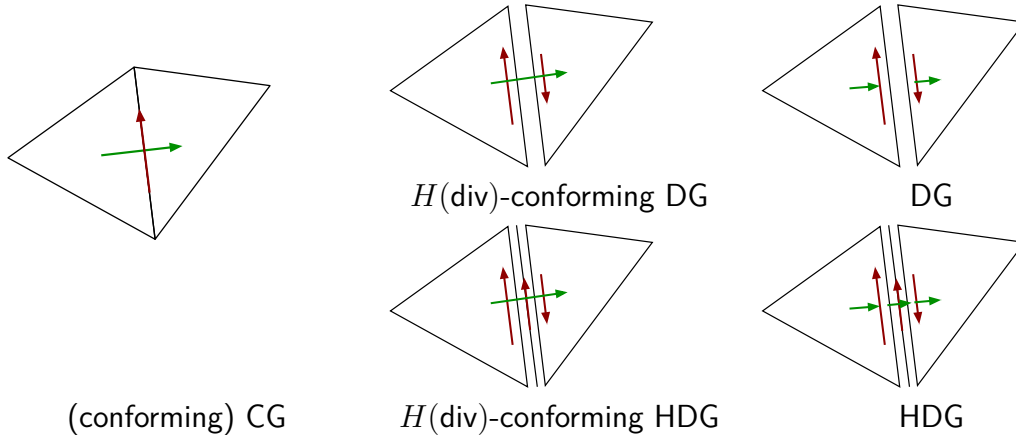

Figure 2.1.1: tangential and normal continuity for different methods

In Figure 2.1.1 a sketch of several possible discretizations of the velocity field is drawn. One very common approach is to consider the conforming continuous Galerkin approach. Here, all components are restricted to be continuous, which is owed to the $[H^1(\Omega)]^d$-regularity of the problem. As a rapid change in velocity and velocity gradients appears in many flows the standard Discontinuous Galerkin approach is also often used. No conformity condition is imposed on the Finite Element space. Of course a hybrid version is most often possible and introduces additional unknowns on the element interfaces for all components.

The approach we are interested in is drawn in the middle. Tangential components are allowed to be discontinuous across element interfaces and are considered for only by means of consistent penalization, i.e. through the formulation. A divergence-free DG version was already proposed in [CKS05]. It uses a local DG formulation. Our approach uses a hybrid version of the (symmetric) interior penalty formulation which was used in [CKS07]. In our formulation only the tangential component of the additional facet functions have to be introduced as normal continuity is enforced on the Finite Element space. The global linear systems that arise after discretization can be reduced to degrees of freedoms associated with the facets, which are:

- normal flow degrees of freedom of the $H(\text{div})$-conforming Finite Element space

- tangential facet degrees of freedom

These are as much degrees of freedom on each facet as we would get for a hybridized version of the completely discontinuous approach. But we still have less overall degrees of freedoms due to the $H(\text{div})$-conforming Finite Element space. The construction of $H(\text{div})$-conforming Finite Elements, which we need for our formulation, is the subject of the next section.

2.2 H(div)-conforming Finite Elements

This section is about $H(\text{div})$-conforming discretizations with Finite Elements. We want to present the concepts and main ideas which are necessary for the comprehension and construction of higher order $H(\text{div})$-conforming Finite Elements.
$H(\text{div})$-conforming Finite Elements are supposed to provide that every discrete function $\underline{u}_h$ of the discretization lies in $H(\text{div}, \Omega)$, that is, that

$$\underline{u}_h \in H(\text{div}, \Omega) = \{\underline{u}_h \in [L^2(\Omega)]^d : \text{div}(\underline{u}_h) \in L^2(\Omega)\} \tag{2.2.1}$$

We work with piecewise polynomials on each element. Thus locally the functions are automatically in $H(\text{div}, T)$. But similar to H^1-conforming Finite Elements we have to enforce another compatibility condition on the global functions to satisfy $\underline{u}_h \in H(\text{div}, \Omega)$. For H^1-conformity the requirement is the continuity over element interfaces, for $H(\text{curl})$-conformity[1] it would be the continuity of the tangential component over element interfaces and in our case, we need continuity of the normal-component.

Lemma 2.2.1. *Let $\mathcal{T}_h$ be a triangulation of a domain Ω. A function $\underline{v}$ is in $H(\text{div}, \Omega)$ if and only if $\underline{v} \in H(\text{div}, T) \ \forall \ T \in \mathcal{T}_h$ and $[\![\underline{v}]\!]_n = 0$ on all interior facets $E \in \mathcal{F}_h^{int}$.*

We will speak of element contributions to the divergence, which is simply $\text{div}(\cdot)$ on each element T (which may represent sources and sinks) and facet contributions, which come from the normal components on the element boundaries (think of the sum of in- and outflows).
$H(\text{div})$-conforming discretizations were first used in the context of mixed formulations of Poisson's equation, as we have already seen in the last chapter. Typical examples for such spaces are the (lowest order) *Raviart-Thomas* and the (lowest order) *Brezzi Douglas Marini* Elements which we want to discuss in the next section. Afterwards the *Piola transformation* will be discussed, which allows for reducing the problem of construction Finite Element shape functions on arbitrary elements to the same problem on a reference element. Further, we present the exact sequence property of H^1, $H(\text{curl})$, $H(\text{div})$ and L^2 in section 2.2.3 which is the foundation for the chosen realization of higher order $H(\text{div})$-conforming methods of section 2.2.4. We conclude with interpolation operators on an $H(\text{div})$-conforming space.
All results on $H(\text{div})$-conformity will be given without proof as we want to focus on the concepts in this section. Additionally we will focus on the two dimensional case involving triangles only. For extensions to three dimensions and proofs we refer to [Zagl06], [SZ05], [DB05] and [DGS08].

2.2.1 Lowest Order H(div)-conforming Finite Elements

Let's start looking at lowest order realizations of those Finite Elements.

---

[1] $H(\text{curl}, \Omega) := \{\underline{u}_h \in [L^2(\Omega)]^d : \underline{\text{curl}}(\underline{u}_h) \in [L^2(\Omega)]^d\}$

Lowest order Raviart Thomas Finite Elements

The most famous $H(\text{div})$-conforming Finite Element is the lowest order *Raviart Thomas* Finite Element, abbreviated by RT_0. An RT_0 Finite Element has m shape functions, where m is the number of facets (faces (in 3D) or edges (in 2D)). Let's focus on two dimensional, triangular meshes here. Those trial functions have non-zero normal component one on one facet and zero on all others. Accordingly every facet possesses one degree of freedom, which is the total flow over a facet:

$$N_E^0(\underline{\psi}) = \int_E \underline{\psi} \cdot \underline{n} \, d\underline{s} \tag{2.2.2}$$

With the help of the barycentric coordinates λ_i, the trial functions can be easily constructed. The trial function associated with the edge E connecting the nodes (w.l.o.g.) 1 and 2 is

$$\underline{\psi}_E^0(\lambda_1, \lambda_2) = \lambda_2 \, \underline{\text{curl}}(\lambda_1) - \lambda_1 \, \underline{\text{curl}}(\lambda_2) \qquad \text{with } \underline{\text{curl}} = \begin{pmatrix} \partial_{x_2} \\ -\partial_{x_1} \end{pmatrix}$$

Those trial functions' element contribution to the divergence are constant by construction as is the normal flow on each edge. We note further that the local space of the RT_0 Finite Element is

$$\text{RT}_0(T) = \{\underline{v}(\underline{x}) = \underline{a} + \underline{x} \cdot b, \underline{a} \in \mathbb{R}^d, b \in \mathbb{R}\} \tag{2.2.3}$$

and as $\text{div}(\underline{x}) = d = const$ the relation

$$\text{div}(\text{RT}_0(\mathcal{T}_h)) = Q_h^0(\mathcal{T}_h) , \qquad \text{with } \text{RT}_0(\mathcal{T}_h) := \bigoplus_{E \in \mathcal{F}_h} \text{span}\{\underline{\psi}_E^0\} \subset H(\text{div}, \Omega) \tag{2.2.4}$$

holds.

Lowest order Brezzi Douglas Marini Finite Elements

We can enrich the RT_0 element by adding one additional shape function on each facet. We therefore introduce the additional DOF which we mark with a star:

$$N_E^*(\underline{\psi}) = \int_E \underline{\psi} \cdot \underline{n} \, v \, d\underline{s} \quad \text{for } v \in \mathcal{P}^1(E) \cap \mathcal{P}^0(E)^\perp$$
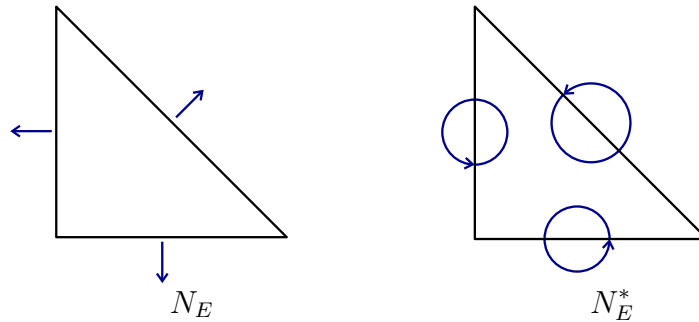

$$N_E \qquad\qquad N_E^*$$

Figure 2.2.1: Degrees of freedom for the lowest order *BDM* element

The associated shape functions are

$$\underline{\psi}_E^*(\lambda_1, \lambda_2) = \underline{\text{curl}}(\lambda_1 \lambda_2)$$

There holds $\mathrm{div}(\underline{\psi}_E^*) = 0$ and $N_E(\underline{\psi}_E^*) = 0 \ \forall E \in \partial T$. The element consisting of both DOF on each facet is called *Brezzi Douglas Marini* element of first order, or short BDM_1. The number of DOF on a simplex exactly matches the dimension of the polynomial space $[\mathcal{P}^1(T)]^d$ and indeed the proposed shape functions are linearly independent and span the local space $[\mathcal{P}^1(T)]^d$.

If we identify the degrees of freedom associated with the faces of the mesh we obtain the global BDM_1 Finite Element Space

$$\mathrm{BDM}_1(\mathcal{T}_h) := \bigoplus_{E \in \mathcal{F}_h} \mathrm{span}\{\underline{\psi}_E^0, \underline{\psi}_E^*\} \subset H(\mathrm{div}, \Omega) \tag{2.2.5}$$

Both elements, RT_0 as well as BDM_1 fulfill

$$\mathrm{div}(\mathrm{RT}_0(\mathcal{T}_h)) = \mathrm{div}(\mathrm{BDM}_1(\mathcal{T}_h)) = Q_h^0(\mathcal{T}_h) \tag{2.2.6}$$

but the approximation of the function itself is better for BDM_1 as

$$[\mathcal{P}^0(T)]^d \subset \mathrm{RT}_0(T) \subset \mathrm{BDM}_1(T) = [\mathcal{P}^1(T)]^d \tag{2.2.7}$$

holds.

For both approaches there exist higher order versions. The local spaces $\mathrm{RT}_k(T)$ and $\mathrm{BDM}_k(T)$ are

$$\mathrm{RT}_k(T) = [\mathcal{P}^k(T)]^d \oplus \underline{x} \cdot \mathcal{P}^{k,*}(T) \tag{2.2.8}$$
$$\mathrm{BDM}_k(T) = [\mathcal{P}^k(T)]^d \tag{2.2.9}$$

where $\mathcal{P}^{k,*}(T)$ are all polynomials on T of exact degree k. So the index k of RT and BDM indicates which polynomial space is completely contained within the respective local spaces. The relations between the higher order spaces is similar to the lower order ones:

$$\mathrm{div}(\mathrm{RT}_k(\mathcal{T}_h)) = \mathrm{div}(\mathrm{BDM}_{k+1}(\mathcal{T}_h)) = Q_h^k(\mathcal{T}_h) \tag{2.2.10}$$

and

$$[\mathcal{P}^k(T)]^d = \mathrm{BDM}_k(T) \subset \mathrm{RT}_k(T) \subset \mathrm{BDM}_{k+1}(T) = [\mathcal{P}^{k+1}(T)]^d \tag{2.2.11}$$

We see that the approximation of the divergence is, w.r.t. the total degrees of freedom, faster achieved with *Raviart Thomas* elements. But if the approximation of the function itself is important the method of choice is the *BDM* element. In our case we use the *BDM* element as we are especially interested in the velocity-field itself. How to choose the functionals N and how to construct appropriate trial functions for the high order *BDM*-case is explained in section 2.2.4. Before we turn over to that, we will show how $H(\mathrm{div})$-conforming Finite Elements on the physical domain can be achieved by means of transforming Finite Elements from the reference domain to the physical domain in a way which preserves the normal component at the boundary and thereby the $H(\mathrm{div})$-conformity.

2.2.2 H(div)-conforming Transformation (Piola Transformation)

To obtain globally $H(\mathrm{div})$-conforming Finite Elements, i.e. shape functions whose normal components at element interfaces coincide also on the physical domain, we make use of the $H(\mathrm{div})$-conforming transformation also known as *Piola* transformation. It further ensures that the introduced degrees of freedom are meaningful also on the physical domain. The following statements on the *Piola* transformation are given without proof. Those can be found in [Zagl06] or [Monk03, section 3.9]

Lemma 2.2.2. *($H(div)$-conforming transformation) Let $\underline{\Phi}_T : \hat{T} \to T$ be a diffeomorphism and $\underline{\hat{u}} \in H(div, \hat{T})$. Then the* Piola transformation

$$\underline{u} := J_T^{-1} \; \underline{\underline{F}}_T \; \underline{\hat{u}} \circ \underline{\Phi}_T^{-1}$$

with $\underline{\underline{F}}_T = \underline{\nabla} \, \underline{\Phi}_T$, $J_T = \det(\underline{\underline{F}}_T)$, implies $\underline{u} \in H(div, T)$. The divergence of $\underline{\hat{u}}$ is transformed even simpler:

$$div_{\underline{x}}(\underline{u}) := J_T^{-1} \; div_{\underline{\hat{x}}}(\underline{\hat{u}} \circ \underline{\Phi}_T^{-1})$$

This transformation results in some further desirable properties:

1. The normal components transform like

$$\underline{u}(\underline{x}) \cdot \underline{n}(\underline{x}) = \underline{\hat{u}}(\underline{\hat{x}}) \cdot \underline{\hat{n}}(\underline{\hat{x}}) \cdot \frac{1}{J_T \|F_T^T \underline{\hat{n}}\|}$$

   The transformation factor cancels out (except for the sign) with the transformation factor of surface integrals

$$\int_E d\underline{s} = \int_{\hat{E}} |J_T| \|F_T^{-T} \underline{\hat{n}}\| \, d\underline{s}$$

   which leads to (except for the sign) invariant surface integrals

$$\int_E \underline{u} \cdot \underline{n}_E \cdot q \; d\underline{s} = \text{sign}(J_T) \int_{\hat{E}} \underline{\hat{u}} \cdot \underline{\hat{n}}_{\hat{E}} \cdot \hat{q} \; d\underline{\hat{s}} \qquad (2.2.12)$$

2. Integrals of the form $\int_T div(\underline{u}) \cdot q \; d\underline{x}$ are invariant (except for the sign) w.r.t. the Piola transformation:

$$\int_T div_{\underline{x}}(\underline{u}) \cdot q \; d\underline{x} = \text{sign}(J_T) \int_T div_{\underline{\hat{x}}}(\underline{\hat{u}}) \cdot \hat{q} \; d\underline{\hat{x}} \qquad (2.2.13)$$

3. In the case of affine linear transformation, the *Raviart Thomas* space $\text{RT}_k(T)$ is invariant with respect to the Piola transformation[2].

With the help of the *Piola transformation* shape functions have to be defined only on the reference domain. Additionally some integral terms are geometry-independent what can be exploited for a faster matrix assembly. Using the *Piola transformation* on curved elements may lead to non-polynomial local spaces on the physical domain as the local space on the reference domain is set to be polynomial and the transformation doesn't have to be. In this case the *Piola transformation* still ensures $H(div)$-conformity as the transformation of the normal component coincides for two (curved) adjacent elements.

2.2.3   The Exact sequence and the DeRham Complex

The exact sequence of the Sobolev spaces H^1, $H(div)$, $H(curl)$ and L^2 characterizes the relationship between spaces and the according differential operators and provides a foundation for the construction of the higher order $H(div)$-conforming Finite Elements we want to use.

---

[2]The same holds trivially for the Brezzi Douglas Marini Element space BDM_k

2.2.3.1   Exact sequence of H^1, $H(\mathbf{div})$, $H(\mathbf{curl})$ and L^2

Let's first state the exact sequence property on the continuous level in two and three dimensions. It is called the *de Rham Complex* and in three dimensions it is written down as

$$\mathbb{R} \xrightarrow{id} H^1(\Omega) \xrightarrow{\nabla} H(\text{curl}, \Omega) \xrightarrow{\text{curl}} H(\text{div}, \Omega) \xrightarrow{\text{div}} L^2(\Omega) \xrightarrow{0} \{0\} \qquad (2.2.14)$$

and should be read in the following way: the kernel of the operator standing on the right side of the space is exactly the image of the operator standing on the left side of the same space. For example the kernel of the <u>curl</u> operator in $H(\text{curl}, \Omega)$ coincides with the gradients of $H^1(\Omega)$.
The exact sequence implies the well known identities $\underline{\text{curl}}(\underline{\nabla}(\cdot)) = 0$ and $\text{div}(\underline{\text{curl}}(\cdot)) = 0$. It actually tells us more: Assume $\text{div}(\underline{u}) = 0$ for a function $\underline{u} \in H(\text{div}, \Omega)$, then there exists a function $\underline{\phi} \in H(\text{curl}, \Omega)$, s.t. $\underline{u} = \underline{\text{curl}}\,\underline{\phi}$ or assume $\underline{\text{curl}}\,\underline{u} = 0$ for a function $\underline{u} \in H(\text{curl}, \Omega)$, then there exists a function $\phi \in H^1(\Omega)$, s.t. $\underline{u} = \underline{\nabla}\phi$.

In two dimensions, the sequence is reduced by one stage, which is either $H(\text{curl})$ and $H(\text{div})$. Thus we have two different versions of the *de Rham Complex*:

$$\mathbb{R} \xrightarrow{id} H^1(\Omega) \xrightarrow{\nabla} H(\text{curl}, \Omega) \xrightarrow{\text{curl}} L^2(\Omega) \xrightarrow{0} \{0\} \qquad (2.2.15a)$$

$$\mathbb{R} \xrightarrow{id} H^1(\Omega) \xrightarrow{\text{curl}} H(\text{div}, \Omega) \xrightarrow{\text{div}} L^2(\Omega) \xrightarrow{0} \{0\} \qquad (2.2.15b)$$

The latter is the more important one for us as it involves the space $H(\text{div}, \Omega)$.

2.2.3.2   The exact sequence for Finite Element spaces

Conforming Finite Elements which also fulfill the exact sequence property are desirable as the separation into the kernel of the according differential operator and the remainder of the space can be exploited. The following diagram includes the finite dimensional analog to the *de Rham Complex* of the continuous spaces and presumes conforming discretizations (we just consider the 2D case (2.2.15b)).

$$
\begin{array}{ccccccccc}
\mathbb{R} & \xrightarrow{id} & H^1(\Omega) & \xrightarrow{\text{curl}} & H(\text{div}, \Omega) & \xrightarrow{\text{div}} & L^2(\Omega) & \xrightarrow{0} & \{0\} \\
 & & \cup & & \cup & & \cup & & \\
\mathbb{R} & \xrightarrow{id} & W_h & \xrightarrow{\text{curl}} & \Sigma_h & \xrightarrow{\text{div}} & Q_h & \xrightarrow{0} & \{0\}
\end{array}
\qquad (2.2.16)
$$

2.2.3.3   Exact sequence for the lowest order Finite Element spaces

The group of Finite Elements consisting of first order nodal Finite Elements, lowest order *Raviart Thomas* Finite Elements and piecewise constant Finite Elements possesses the exact sequence property[3], which is drawn in Figure 2.2.2. The two numbers below the drawn Finite Element indicate the dimension of the image of the left hand side operator which coincides with the dimension of the kernel of the right hand side operator (left number) and the dimension of the rest of the Finite Element space (right number). The sum of both is the number of degrees of freedom on each element.

---

[3]in 3D, the lowest order Nedelec Finite Elements of first kind would be the appropriate representative for $H(\text{curl})$ Finite Elements
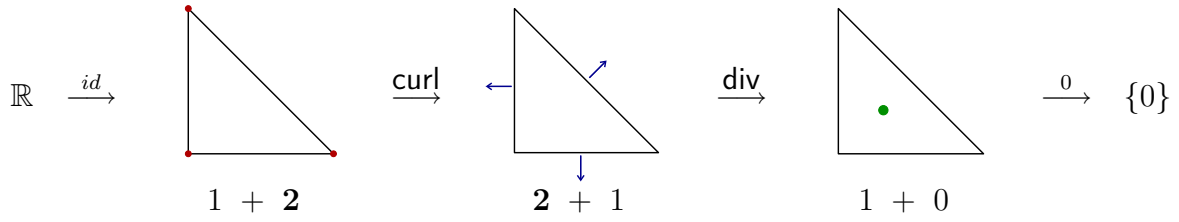
Figure 2.2.2: exact sequence involving the lowest order *Raviart Thomas* element

Another exact sequence involves the BDM$_1$ element and is the following.
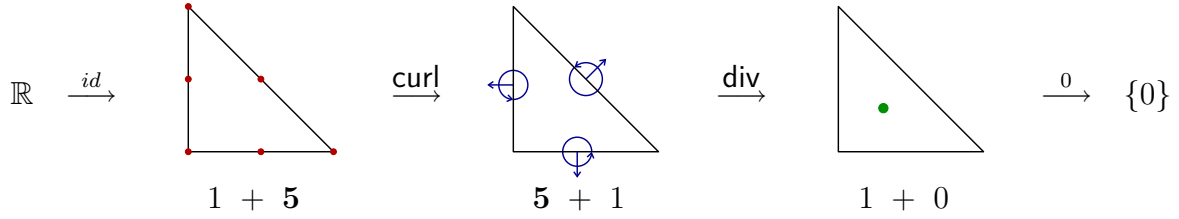
Figure 2.2.3: exact sequence involving the lowest order *BDM* element

We see that the additional edge DOF of the H^1-conforming element lead to three additional DOF for the $H(\text{div})$-conforming Finite Element space. But as we already saw, possible additional shape functions are $\underline{\psi}^*_E(\lambda_1, \lambda_2) = \underline{\text{curl}}(\lambda_1\lambda_2)$ which are $\underline{\text{curl}}$-fields (of H^1-functions)[4] and thereby divergence-free. The last step of the sequence is thus uneffected. This exact sequence reduces the polynomial order by one for each differential operator in contrast to the one involving the *Raviart Thomas* element.

2.2.4   Construction of Higher Order H(div)-conforming Finite Elements

After having discussed the lowest order cases, we want to turn to the construction of $H(\text{div})$-conforming Finite Elements with arbitrary polynomial degree > 1. Several realizations are possible. Our approach of choice is the one proposed and extensively discussed in [SZ05, Zagl06], where the construction of variable order H^1-, $H(\text{curl})$-, $H(\text{div})$- and L^2-conforming Finite Element spaces on hybrid meshes with tetrahedrals, hexahedrals, prisms, triangles and quadrilaterals is presented. The construction of the higher order shape functions mimics the exact sequence property of the spaces H^1, $H(\text{curl})$, $H(\text{div})$ and L^2 and thereby facilitates further approvements.

We proceed as follows: For the $H(\text{div})$-conforming case only, we will show which face- and cell degrees of freedom should be used to enable an $H(\text{div})$-conforming Finite Element space which inherits the exact sequence property from the continuous level. For completeness we will also give the set of basis functions which were derived in [SZ05, Zagl06]. These use the separation of the element-local Finite Element space into low order RT$_0$, face-associated, divergence-free element-associated and polynomial space completing basis functions. This classification allows for reducing the degrees of freedom without loss of accuracy, which will be discussed in the last part of this section. As before, we will limit our discussion on the concepts and the two-dimensional case with a p-uniform triangular meshes. For the more involved cases and further details we recommend to consult [Zagl06], where all details are included and comprehensibly explained.

---

[4]$\lambda_1\lambda_2$ is actually the standard second order edge bubble function

2.2.4.1 Transfering the exact sequence property to the discrete level

We now want to obtain a sequence of higher order Finite Element spaces (W_h, Σ_h, Q_h) which fulfill (2.2.16). All Finite Element spaces should be approximated with piecewise complete polynomials, which is the generalization of the lowest order case of the exact sequence which involves the BDM_1 element. We devide the degrees of freedom into three classes:

1. lowest order degrees of freedom (vertex-, edge- and cell-based[5])

2. higher order edge[6] degrees of freedom

3. higher order element degrees of freedom

The lowest order degrees of freedom are exactly those we already used in Figure 2.2.2, i.e. vertex shape functions, RT_0 and piecewise constants. For each polynomial degree an additional degree of freedom is associated with an edge for W_h and Σ_h. These are called higher order edge degrees of freedom and ensure continuity and normal-continuity, respectively. All other degrees of freedom are higher order element degrees of freedom and have to provide shape functions which are zero at the element boundary for W_h or which have zero normal component for Σ_h. For Q_h no such classification has to be made as no continuity restriction is necessary.

In the next section we will give an appropriate choice of moments for $H(\text{div})$-conforming Finite Elements only, before we discuss the separation of the Finite Element space into four classes.

Moments / Degrees of freedom

We already discussed the lowest order degrees of freedom related to the RT_0 and BDM_1 Finite Elements in section 2.2.1. In the high order case we use the following degrees of freedom for a Finite Element of uniform order k:

- lowest order edge-based DOF:
$$N_E(\underline{\phi}) = \int_E \underline{\phi} \cdot \underline{n} \, d\underline{s}$$

- higher order edge-based DOF:
$$N_E^*(\underline{\phi}) = \int_E \underline{\phi} \cdot \underline{n} \, v \, d\underline{s} \quad \text{for } v \in \mathcal{P}^k(E) \cap \mathcal{P}^0(E)^\perp$$

- higher order nonzero divergence cell-based DOF:
$$N_C^A(\underline{\phi}) = \int_T \underline{\phi} \cdot \underline{v}_i \, d\underline{x} \quad \text{for } \{\underline{v}_i\} \text{ a basis of curl}([\mathcal{P}^{k+1}]_{t,0}^d)$$

where $[\mathcal{P}^k]_{t,0}^d = [\mathcal{P}^k]^d \cap H_0(\text{curl}, T)$

- higher order divergence-free cell-based DOF:
$$N_C^B(\underline{\phi}) = \int_T \text{div}(\underline{\phi}) \, \text{div}(\underline{v}_i) \, d\underline{x} \quad \text{for } \{\underline{v}_i\} \text{ s.t } \{\text{div}(\underline{v}_i)\} \text{ is a basis of div}([\mathcal{P}^k]_{n,0}^d)$$

where $[\mathcal{P}^k]_{n,0}^d = [\mathcal{P}^k]^d \cap H_0(\text{div}, T)$

---

[5] in 3D vertex-, face-, edge- and cell-based
[6] in 3D face and edge

The last separation of cell-based DOF is owed to the exact sequence as it naturally divides them into shape functions which have nonzero divergence ($\leftrightarrow\ N_C^B$) and those who are divergence-free ($\leftrightarrow\ N_C^A$). Additionally the higher order edge-based shape functions can be taken as the curl of H^1 edge-based shape functions and are thereby also divergence-free.

We consider the following shape functions related to the degrees of freedom we just defined:

1. edge-based shape functions for edge E connecting vertices (w.l.o.g.) 1 and 2:

   (a) lowest order:
   $$\underline{\psi}_E^0 = \underline{\mathrm{curl}}(\lambda_1)\lambda_2 - \underline{\mathrm{curl}}(\lambda_2)\lambda_1$$

   (b) higher order:
   $$\underline{\psi}_E^i = \underline{\mathrm{curl}}(L_{i+1}^{\mathcal{S}}(\lambda_1 - \lambda_2, \lambda_1 + \lambda_2))\ \ [= \underline{\mathrm{curl}}(\varphi_E^{i+1})], \qquad \text{for } 1 \leq i \leq k$$

2. cell-based shape functions for element T sharing vertices 1, 2 and 3:

   (a) divergence-free:
   $$\underline{\psi}_C^{A;i,j} = \underline{\mathrm{curl}}(u_i)v_j - \underline{\mathrm{curl}}(v_j)u_i = \underline{\mathrm{curl}}(u_i\ v_j)\ \ [= \underline{\mathrm{curl}}(\varphi_C^{i,j})] \qquad \text{for } 0 \leq i + j \leq k - 2$$

   (b) nonzero divergence:
   $$
   \begin{array}{rll}
   \underline{\psi}_C^{B_1;i,j} & = & \underline{\mathrm{curl}}(u_i)v_j + \underline{\mathrm{curl}}(v_j)u_i \qquad\qquad \text{for } 0 \leq i+j \leq k-2 \\
   \underline{\psi}_C^{B_2;j} & = & (\underline{\mathrm{curl}}(\lambda_1)\lambda_2 - \underline{\mathrm{curl}}(\lambda_2)\lambda_1)v_j \qquad \text{for } 0 \leq j \leq k-2
   \end{array} \qquad (2.2.17)
   $$

   where $u_i := L_{i+2}^{\mathcal{S}}(\lambda_1 - \lambda_2, \lambda_1 + \lambda_2)$ and $v_j := \lambda_3 l_j(2\lambda_3 - 1)$. Here we made use of the *Scaled Integrated Legendre Polynomials* of degree $i+2$ $L_{i+2}^{\mathcal{S}}$ and the *Legendre polynomials* l_j of degree j. Both are defined in the appendix A.3.

The functions φ_E^{i+1} and $\varphi_C^{i,j}$ are the edge- and cell-based shape functions used in [SZ05, Zagl06] for the H^1-conforming Finite Element space. So we see $\underline{\mathrm{curl}}(W_h)$ is completely included in Σ_h and furthermore the subspace $\underline{\mathrm{curl}}(W_h \setminus \{\phi_V\})$, where $\{\phi_V\}$, the set of all lowest order vertex-based shape functions, is explicitly separable, i.e. we can, except for the lowest order RT$_0$ functions, divide the space into divergence and divergence-free functions.

Before we will make use of this, let's summarize the properties of the separation of the Finite Element space in the next Table (2.1). Note that now each symbol (e.g. an arrow) corresponds *not* to exactly one degree of freedom but to a class of degrees of freedom.
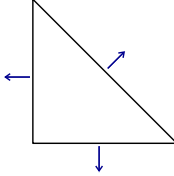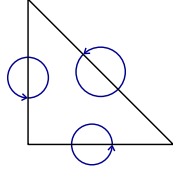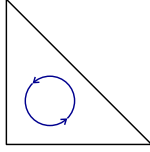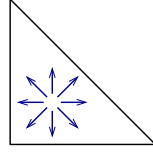
RT$_0$ DOF	higher order edge DOF	higher order div.-free DOF bubble	higher order DOF with nonz. div

$$\Sigma_h^k = \bigoplus_{E\in\mathcal{F}_h} \mathrm{span}\{\underline{\psi}_E^0\} \;\oplus\; \bigoplus_{E\in\mathcal{F}_h} \mathrm{span}\{\underline{\psi}_E^i\} \;\oplus\; \bigoplus_{T\in\mathcal{T}_h} \mathrm{span}\{\underline{\psi}_C^{A;i,j}\} \;\oplus\; \bigoplus_{T\in\mathcal{T}_h} \mathrm{span}\{\underline{\psi}_C^{B_1;i,j}, \underline{\psi}_C^{B_2;j}\}$$

$$\mathrm{div}(\Sigma_h^k) = \bigoplus_{T\in\mathcal{T}_h} \mathcal{P}^0(T) \;\oplus\; \{0\} \;\oplus\; \{0\} \;\oplus\; \bigoplus_{T\in\mathcal{T}_h} [\mathcal{P}^{k-1}\cap\mathcal{P}^{0\perp}](T)$$

$$\#\mathrm{DOF} = 3 \;+\; 3k \;+\; \tfrac{1}{2}k(k-1) \;+\; \tfrac{1}{2}k(k+1) - 1$$

Table 2.1: Separation of the $H(\mathrm{div})$-conforming Finite Element space

Let's also try a more physical interpretation of this space separation. Every flow field can be interpreted as a superposition of a divergence- and vortex-free mean flow, (smaller and larger) eddies, sources and sinks. The eddies can again be subdivided into three kind of eddies:

1. eddies aroung one or more vertices

2. eddies on an edge, i.e. eddies which have zero normal velocity on all edges except for one

3. eddies which lie within one element

To resolve the mean flow and the first kind of eddies we need the lowest order RT$_0$ shape functions. Now the eddies we need to represent in order to increase the resolution can either be located at one edge or within an element.

2.2.4.2 Excluding higher order divergence functions - Reducing the basis

With the introduced separation, we can neglect all higher order nonzero divergence shape functions if we consider an incompressible flow explicitly. For $\underline{u} \in \Sigma_h^k$ there holds

- the divergence of the higher order divergence shape functions are linear independent (due to the choice of the functionals N)

- the divergence of the space of all higher order divergence shape functions is $\frac{1}{2}p(p+1) - 1$ dimensional on each element

- the divergence of the higher order divergence shape functions is L^2-orthogonal to (element-piecewise) constant functions (due to $\int_{\partial T} \underline{\psi}_C^B \cdot \underline{n}\,ds = \int_T \mathrm{div}(\underline{\psi}_C^B) = 0$)

- we can uniquely decompose $\underline{u}$ into $\underline{u}_E^{\mathrm{RT0}}$, $\underline{u}_E^{ho}$, $\underline{u}_C^A$ and $\underline{u}_C^B$ due to the separation of the space.

- $\mathrm{div}(\underline{u}) = 0$ now requires $\mathrm{div}(\underline{u}_E^{\mathrm{RT0}}) = 0$, where $\mathrm{div}(\underline{u}_E^{\mathrm{RT0}}) \in Q_h^0(\mathcal{T}_h)$ and $\mathrm{div}(\underline{u}^B) = 0$, where $\mathrm{div}(\underline{u}^B) \in \mathcal{P}^{k-1}(T) \cap \mathcal{P}^0(T)^\perp$ on each element T

- This gives $\frac{1}{2}p(p+1) - 1$ equations for $\underline{u}^B$ which is a linear combination of $\frac{1}{2}p(p+1) - 1$ shape functions.

Thus $\mathrm{div}(\underline{u}) = 0$ always implies $\underline{u}^B = 0$, s.t. those DOF can be neglected.

This reduces the number of unknowns drastically. Additionally for the incompressible Navier-Stokes equations the pressure unknowns, which are lagrange multipliers for the incompressibility constraint, can be reduced to the lowest order case, i.e. piecewise constant functions, because we only have to control the RT_0 degrees of freedom. The overall unknowns can thereby approximately halfed without loss of accuracy for higher order approximation ($p \geq 2$, see section 2.5.2). The neglected velocity DOF would be zero the one way or the other. Only the pressure approximation has to be recovered with a post-processing step. This is not discussed in the proceeding. So instead of Σ_h^k we can work with the space

$$\Sigma_h^{k,*} := \bigoplus_{E \in \mathcal{F}_h} \mathrm{span}\{\psi_E^0\} \oplus \bigoplus_{E \in \mathcal{F}_h} \mathrm{span}\{\psi_E^i\} \oplus \bigoplus_{T \in \mathcal{T}_h} \mathrm{span}\{\psi_C^A\} \tag{2.2.18}$$

2.2.5   H(div)-conforming Interpolation

At several places during the analysis later on we will need an $H(\mathrm{div})$-conforming interpolation $\underline{\underline{\Pi}}_{h,k}^\Sigma :$ $[H^1]^d \rightarrow \Sigma_h^k$, which possesses the following properties:

- $H(\mathrm{div})$-conformity: $\underline{\underline{\Pi}}_k^\Sigma(\underline{u}) \in \Sigma_h^k$

- L^2-optimal approximation of the divergence: $(\mathrm{div}(\underline{\underline{\Pi}}_k^\Sigma(\underline{u})), q_h) = (\mathrm{div}(\underline{u}), q_h) \quad \forall\, q_h \in Q_h^{k-1}$

- boundedness of the interpolation in the H^1 norm: $\|\underline{\underline{\Pi}}_k^\Sigma(\underline{u})\|_{H^1} \leq C\|(\underline{u})\|_{H^1}$

- the interpolator is a projector, s.t. $\underline{\underline{\Pi}}_k^\Sigma(\underline{v}) = \underline{v} \qquad \forall\, \underline{v} \in \Sigma_h^k$

Those properties can be achieved with the interpolation operator proposed in [BF91, III.3.3]. Also the projection-based interpolation proposed in [DB05] fulfills all the above requirements and is furthermore p-robust. Proofs and further discussion can be found in [DB05] and [DGS08].

2.3 Exactly divergence-free Hybrid DG method for the Stokes problem

In this section we want to present the variational formulation as well the a priori error analysis for the steady Stokes problem, which reads as

$$
\begin{cases}
\operatorname{div}(-\nu \underline{\nabla}\, \underline{u} + p\underline{I}) &= \underline{f} & \text{in } \Omega \\
\operatorname{div}(\underline{u}) &= 0 & \text{in } \Omega \\
\underline{u} &= \underline{u}_D & \text{on } \Gamma_D \\
(\nu \underline{\nabla}\, \underline{u} - p\underline{I}) \cdot \underline{n} &= 0 & \text{on } \Gamma_{out}
\end{cases}
\tag{2.3.1}
$$

On Γ_{out} the natural boundary conditions are used. Those boundaries prescribe zero momentum flux on the outflow region. It naturally comes up like Neumann boundaries ($\frac{\partial u}{\partial \underline{n}}[= 0]$) in the scalar case.

2.3.1 Introducing the method

As we already pointed out in section 2.1 our approach to a good (H)DG method is to decompose the H^1-consistency into normal and tangential consistency. Both components will be treated differently. The normal consistency (i.e. normal continuity) will be enforced by the choice of the discrete space. The tangential consistency will be included with a HDG formulation. The procedure to reduce trial and test functions will not be adressed explicitly as the results stay the same and only make a difference for the implementation.

viscous part

The procedure of deriving our Hybrid DG formulation for the viscous part is similar to the one used for scalar diffusion. First we integrate by parts on each element.

$$
\int_\Omega \operatorname{div}(-\nu \underline{\nabla}\, \underline{u})\underline{v}\, d\underline{x} = \sum_{T \in \mathcal{T}_h} \int_T \operatorname{div}(-\nu \underline{\nabla}\, \underline{u})\underline{v}\, d\underline{x} = \sum_{T \in \mathcal{T}_h} \int_T \nu \underline{\nabla}\, \underline{u} : \underline{\nabla}\, \underline{v}\, d\underline{x} - \int_{\partial T} \nu \frac{\partial \underline{u}}{\partial \underline{n}} \cdot \underline{v}\, d\underline{s} \tag{2.3.2}
$$

Adding a consistent term

$$
\sum_{T \in \mathcal{T}_h} \int_{\partial T} \nu \frac{\partial \underline{u}}{\partial \underline{n}} \cdot \underline{\tilde{v}}\, d\underline{s} = \int_{\Gamma_N} \nu \frac{\partial \underline{u}}{\partial \underline{n}} \cdot \underline{\tilde{v}}\, d\underline{s} \tag{2.3.3}
$$

with the function $\underline{\tilde{v}} = \underline{v}^n + \underline{v}_F^t$, which is single-valued, results in

$$
\sum_{T \in \mathcal{T}_h} \int_T \nu \underline{\nabla}\, \underline{u} : \underline{\nabla}\, \underline{v}\, d\underline{x} - \int_{\partial T} \nu \frac{\partial \underline{u}}{\partial \underline{n}} \cdot \underbrace{(\underline{v} - \underline{\tilde{v}})}_{=[\![\mathbf{v}^t]\!]}\, d\underline{s} - \int_{\Gamma_N} \nu \frac{\partial \underline{u}}{\partial \underline{n}} \cdot \underline{\tilde{v}}\, d\underline{s} \tag{2.3.4}
$$

where we used $\underline{v} - \underline{\tilde{v}} = \underline{v}^t - \underline{v}_F^t =: [\![\mathbf{v}^t]\!]$. The boundary integral allows for naturally plugging in the outflow boundary condition. On the other parts the boundary term will vanish as $\tilde{v}$ is zero there. Similar to the procedure in the scalar case we can now use that $[\![\mathbf{u}^t]\!]$ is zero for the true solution to add consistency terms for symmetry and stability. Finally we end up with

$$
\mathcal{B}_h(\underline{u}, \underline{v}) := \sum_{T \in \mathcal{T}_h} \int_T \nu \underline{\nabla}\, \underline{u} : \underline{\nabla}\, \underline{v}\, d\underline{x} - \int_{\partial T} \nu \frac{\partial \underline{u}}{\partial \underline{n}} \cdot [\![\mathbf{v}^t]\!]\, d\underline{s} - \int_{\partial T} \nu \frac{\partial \underline{v}}{\partial \underline{n}} \cdot [\![\mathbf{u}^t]\!]\, d\underline{s} + \int_{\partial T} \nu\, \tau_h\, [\![\mathbf{u}^t]\!] \cdot [\![\mathbf{v}^t]\!]\, d\underline{s} \tag{2.3.5}
$$

This actually is the analog to the scalar case. Instead of $[\![\underline{\mathbf{u}}^t]\!]$ and $[\![\mathbf{v}^t]\!]$ you could also write $[\![\underline{\mathbf{u}}]\!]$ and $[\![\mathbf{v}]\!]$ as $[\![\underline{\mathbf{u}}^n]\!]$ and $[\![\mathbf{v}^n]\!]$ is always zero due to $H(\operatorname{div})$-conformity. However, we stick with $[\![\underline{\mathbf{u}}^t]\!]$ and $[\![\mathbf{v}^t]\!]$ to highlight the decomposition of the continuity.

pressure part

For the pressure we again integrate by parts on each element and replace the boundary value for p with a numerical flux $\hat{p}$.

$$\int_{\Omega} \underline{\nabla} p \cdot \underline{v} \, d\underline{x} = \sum_{T \in \mathcal{T}_h} \int_T \underline{\nabla} p \cdot \underline{v} \, d\underline{x} = \sum_{T \in \mathcal{T}_h} - \int_T p \, \mathrm{div}(\underline{v}) \, d\underline{x} + \int_{\partial T} \hat{p} \, v_n \, d\underline{s} \qquad (2.3.6)$$

We choose $\hat{p} = \begin{cases} \{\!\{p\}\!\} & \text{on inner facet} \\ p & \text{on boundary facets} \end{cases}$, s.t. $\displaystyle\sum_{T \in \mathcal{T}_h} \int_{\partial T} \hat{p} \, v_n \, d\underline{s} = \int_{\Gamma_{out}} p \, v_n \, d\underline{s}$ and results in

$$\sum_{T \in \mathcal{T}_h} - \int_T p \, \mathrm{div}(\underline{v}) \, d\underline{x} + \int_{\Gamma_{out}} p \, v_n \, d\underline{s} \qquad (2.3.7)$$

The boundary integrals are part of the natural boundary condition and so won't be considered for in the bilinearform:

$$\mathcal{D}_h(\mathbf{v}, p) := \sum_{T \in \mathcal{T}_h} - \int_T p \, \mathrm{div}(\underline{v}) \, d\underline{x} \qquad (2.3.8)$$

The discrete problem

Now we have described the formulation, but didn't specify completely which spaces we use. First we have to introduce a new facet Finite Element space for the tangential facet functions:

$$F_h^k := \left\{ \underline{u}_F^t : \underline{u}_F^t \in [\mathcal{P}^k(E)]^d \times \underline{n} \; \forall \, E \in \mathcal{F}_h \right\} \qquad (2.3.9)$$

Then our velocity approximation $\underline{\mathbf{u}}$ is again a composition of an element function $\underline{u} \in \Sigma_h^k$ and $\underline{u}_F \in F_h^k$ which we again indicate by the bold writing.

The compound space is called $S_h^k := \Sigma_h^k \times F_h^k$ and $S_{h,D}^k = \Sigma_{h,D}^k \times F_{h,D}^k$ (or $D = 0$) denotes that (homogeneous) dirichlet boundary conditions are incorporated into the space. Normal components of boundary values have to be imposed on Σ_h^k and tangential ones on F_h^k. That means that the normal component is imposed in a strong sense, whereas the tangential component is only enforced in a weak (Hybrid) DG sense.

Equally to the scalar case, the pressure space is the space Q_h^k which consists of discontinuous, piecewise polynomials of order k. It is chosen to be one order less than the velocity approximation.

Then the discrete problem reads:

Find $\underline{\mathbf{u}} \in S_{h,D}^{k+1}$ and $p \in Q_h^k$, s.t.

$$\begin{cases} \mathcal{B}_h(\underline{\mathbf{u}}, \underline{\mathbf{v}}) + \mathcal{D}_h(\underline{\mathbf{v}}, p) &= \langle \underline{f}, v \rangle \quad \forall \, v \in S_{h,0}^{k+1} \\ \mathcal{D}_h(\underline{\mathbf{u}}, q) &= \quad\;\; 0 \quad \forall \, q \in Q_h^k \end{cases} \qquad (2.3.10)$$

We also introduce the larger Bilinearform to shorten the notation at some parts of the analysis:

$$\mathcal{K}_h\left((\underline{\mathbf{u}}, p), (\underline{\mathbf{v}}, q)\right) := \mathcal{B}_h(\underline{\mathbf{u}}, \underline{v}) + \mathcal{D}_h(\underline{\mathbf{v}}, p) + \mathcal{D}_h(\underline{\mathbf{u}}, q) \qquad (2.3.11)$$

Thus (2.3.10) can also be written as (with $U = (\underline{\mathbf{u}}, p)$ the velocity-pressure pair)

$$\mathcal{K}_h(U, V) = \langle f, v \rangle \quad \forall \, V \in S_h^{k+1} \times Q_h^k =: Z_h^k \qquad (2.3.12)$$

We additionally introduce an appropriate norm:

$$\|U\|_{\mathcal{K}_h} = \|(\mathbf{u}, p)\|_{\mathcal{K}_h} := \sqrt{\nu}\|\mathbf{u}\|_1 + \|p\|_{L^2} \tag{2.3.13}$$

where $\|\cdot\|_1$ is the hybrid version of the modified broken H^1-norm, similar to the one we saw during the analysis of the scalar poisson problem:

$$\|\mathbf{u}\|_1^2 := \sum_{T \in \mathcal{T}_h} \left\{ \|\nabla\, \underline{u}\|_T^2 + \frac{1}{h}\|[\![\underline{u}^t]\!]\|_{\partial T}^2 + h\|\frac{\partial u}{\partial \underline{n}}\|_{\partial T}^2 \right\} \tag{2.3.14}$$

Remark 2.3.1 (Weak incompressibility $\Rightarrow$ strong incompressibility):
Due to the choice of the velocity-pressure pair, there holds $\text{div}(\Sigma_h^{k+1}) \subseteq Q_h^k$. Then $\mathcal{D}_h(\mathbf{u}, q) = 0 \; \forall \; q$ results in $\text{div}(\underline{u}) = 0$ in a strong sense on each element. Additionally taking normal continuity, i.e. $H(\text{div})$-conformity, into account gives $\text{div}(\underline{u}) = 0$ globally, thus the approximate solution is actually exactly divergence-free. In other words the formulation imposes a strong incompressibility condition on $\underline{\mathbf{u}}$.

Remark 2.3.2 (Reducing the facet unknowns):
As was pointed out in Remark 1.2.4 for the scalar case, in the pure diffusive limit the jump operator could be replaced by a projection on the polynomial space of a degree lower and so the facet functions could be chosen of an order less. As soon as we also consider convection this is not possible without reducing the convergence rate.

Interpolation on the hybrid space S_h
Owed to the fact that tangential facet functions where added to Σ_h^k, resulting in S_h^k, we have to introduce an appropriate interpolator / projector for the complete Finite Element space S_h:

$$\mathbf{\Pi}_{h,k}^S(\underline{u}) := \left(\Pi_{h,k}^\Sigma(\underline{u}),\; \Pi_{h,k}^{f,t}(\underline{u})\right) \quad \text{with} \quad \underline{\Pi}_{h,k}^{f,t}(\underline{u}) \in F_h^{t,k} \quad \text{and}$$

$$\left(\underline{\Pi}_{h,k}^{f,t}(\underline{u}), \underline{v}_F^t\right)_{\partial T} = \left(\text{tr}|_{\mathcal{F}_h}(\underline{u}), \underline{v}_F^t\right)_{\partial T} \qquad \forall\; \underline{v}_F^t \in F_h^{t^k}(T) \tag{2.3.15}$$

Notice that both projections $\underline{w} = \Pi_{h,k}^\Sigma(\underline{u})$, $\underline{w}_F^t = \Pi_{h,k}^{f,t}(\underline{u})$ applied on adjacent elements result in the same facet function $\tilde{\underline{w}} = \underline{w}^n + \underline{w}_F^t$, so the global projector can be constructed in an element by element fashion. Furthermore the projection $\mathbf{\Pi}_{h,k}^S(\underline{u})$ is bounded in the $\|\cdot\|_1$ norm as $\underline{\Pi}_{h,k}^\Sigma(\underline{u})$ and $\underline{\Pi}_{h,k}^{f,t}(\underline{u})$ are bounded in weaker norms. Of course, as $\mathbf{\Pi}_{h,k}^S(\underline{u})$ is actually a projection onto S_h^k there holds $\mathbf{\Pi}_{h,k}^S(\underline{u}_h) = \underline{u}_h \; \forall \; \underline{u}_h \in S_h^k$.
Let's state the following Lemma, which will be made use of shortly after:

Lemma 2.3.1. *The interpolation operator* $\mathbf{\Pi}_{h,k}^S : [H^1(\Omega)]^d \rightarrow S_h^k$ *for* $k \geq 1$ *fulfills the following relations with* c *independent of the mesh size:*

$$\mathcal{D}_h(\mathbf{u}, q) = \mathcal{D}_h(\mathbf{\Pi}_{h,k}^S \underline{u}, q) \qquad \forall\; q \in Q_h^{k-1} \tag{2.3.16}$$

$$\|\mathbf{\Pi}_{h,k}^S \underline{u}\|_1 \leq c\|\underline{u}\|_{H^1} \tag{2.3.17}$$

Proof. See the appendix, chapter B. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

2.3.2   A priori error analysis

To show well-posedness and later on optimal convergence of the method we will make use of Brezzi's Theorem. Essential conditions for this theorem are the coercivity of $\mathcal{B}_h$, the boundedness of both bilinearforms $\mathcal{B}_h$ and $\mathcal{D}_h$ and an inf-sup condition on $\mathcal{D}_h$.

For the continuous problem with the canonical bilinearforms, i.e. replacing $\mathcal{B}_h$ with $\int_\Omega \nabla \, \underline{u} \, \nabla \, \underline{v} \, d\underline{x}$, coercivity, boundedness and inf-sup condition hold, and due to the subsequent Theorem the problem is well-posed and thus has a unique solution $\underline{u} \in [H^1(\Omega)]^d$ and $p \in L^2(\Omega)$. The task now is to show that the same properties translate to our discretization. Therefore many of the scalar results can be overtaken for Stokes.

The boundary is splitted into two parts Γ_D and Γ_{out} where either Dirichlet boundary conditions or Neumann boundary conditions are prescribed. By Neumann boundary conditions we mean boundaries where $\left(\nu \nabla \, \underline{u} - p\underline{I} \right) \cdot \underline{n}$ is given. We call this part of the boundary Γ_{out} as we just consider homogeneous Neumann boundary conditions which are normally the outflow conditions of a domain. Further we assume that wherever Dirichlet boundary conditions are prescribed, they are prescribed for all components (just for ease of notation) and that Γ_D is a nonempty part of the boundary.

We will use the following discrete and continuous spaces during the analysis:

$$
\begin{aligned}
\mathrm{BDM}_k &= \Sigma_h^k &\subset& \quad \Sigma &=& H(\mathrm{div}, \Omega) \\
\left\{ \underline{u}_F^t : \underline{u}_F^t \in [\mathcal{P}^k(E)]^d \times \underline{n} \ \forall \ E \in \mathcal{F}_h \right\} &= F_h^{t,k} &\subset& \quad F^t &=& \left\{ \underline{u}_F^t : \underline{u}_F^t \in [L^2(\mathcal{F}_h)]^d \times \underline{n} \right\} \\
\Sigma_h^k \times F_h^{t,k} &= S_h^k &\subset& \quad S &=& \Sigma \times F^t \qquad (2.3.18) \\
\left\{ p : p \in \mathcal{P}^k(T) \ \forall \ T \in \mathcal{T}_h \right\} &= Q_h^k &\subset& \quad Q &=& L^2(\Omega) \\
S_h^k \times Q_h^{k-1} &= Z_h^k &\subset& \quad Z &=& [H^1(\Omega)]^d \cap [H^2(\mathcal{T}_h)]^d \times L^2(\Omega)
\end{aligned}
$$

We introduce further the spaces $\Sigma_h^{*,k}$ and $S_h^{*,k}$ which in contrast to $\Sigma_h^{k,*}$ (which was the Finite Element space without the higher order nonzero divergence functions) indicate the divergence-free subspaces of Σ_h^k and S_h^k:

$$
S_h^{*,k} := \{\mathbf{u} \in S_h^k : \mathrm{div}(\underline{u}) = 0\} \overset{Rem.2.3.1}{\Longleftrightarrow} \{\mathbf{u} \in S_h^k : \mathcal{D}_h(\mathbf{u}, q) = 0 \ \forall \ q \in Q_h^{k-1}\} \qquad (2.3.19)
$$

As we need to mimic H^1 regularity the *BDM* approach fits better than *Raviart Thomas* element as with the full polynomial space the approximation of $\nabla \, \underline{u}$ and $\underline{u}$ is more natural.

The identity operator $\underline{\mathbf{id}}_h : [H^1(\Omega)]^d \cap [H^2(\mathcal{T}_h)]^d \to S$, which uses the well-defined tangential trace for the tangential facet functions, will again be used at several occasions implicitely. It provides an extension of the bilinearforms, s.t. $\mathcal{B}_h : S \times S \to \mathbb{R}$ and $\mathcal{D}_h : Q \times S \to \mathbb{R}$ are well defined.

For ease of presentation we only consider constant viscosity ν.

Brezzi's Theorem
We restate a rather extensive version of Brezzi's Theorem:

Theorem 2.3.2 (Brezzi's Theorem)**.** *Assume that we have two Hilbert spaces V and Q and two continuous bilinearforms $a: V \times V \to \mathbb{R}$ and $b: Q \times V \to \mathbb{R}$*

$$
\begin{aligned}
a(u,v) &\leq \beta_a \|u\|_V \|v\|_V \quad \forall\, u,v \in V \\
b(v,q) &\leq \beta_b \|v\|_V \|q\|_Q \quad \forall\, u \in V,\ \forall\, q \in Q
\end{aligned}
$$

and associated linearforms $f: V \to \mathbb{R}$ and $g: Q \to \mathbb{R}$, which are bounded in the operator norms $\|\cdot\|_{V'}$ and $\|\cdot\|_{Q'}$. Assume further that coercivity of $a(\cdot,\cdot)$ on the kernel of b, i.e.

$$
a(u,u) \geq \alpha_a \|u\|_V^2 \quad \forall\, u \in V^* := \{v \in V : b(v,q) = 0 \ \forall\, q \in Q\}
$$

and the inf-sup condition holds

$$
\inf_{0 \neq q \in Q} \sup_{u \in V} \frac{b(u,q)}{\|u\|_V \|q\|_Q} \geq \alpha_b > 0 \quad \Leftrightarrow \quad \sup_{u \in V} \frac{b(u,q)}{\|u\|_V} \geq \alpha_b \|q\|_Q \qquad \forall\, q \in Q
$$

Then, the compound bilinearform

$$
k\left((u,p),(v,q)\right) := a(u,v) + b(u,q) + b(v,p)
$$

is inf-sup-stable, i.e. with the definition of the compound functions, spaces and norms $W := V \times Q$, $U := (u,p)$, $V = (v,q)$ and $\|U\|_W := \|u\|_V + \|p\|_Q$ there holds:

$$
\sup_{V \in W} \frac{k(U,V)}{\|V\|_W} \geq \alpha_k \|U\|_W \qquad \forall\, U \in W
$$

So the mixed problem

$$
\begin{cases}
a(u,v) + b(v,p) &= f(v) \quad \forall\, v \in V \\
b(u,q) &= g(q) \quad \forall\, q \in Q
\end{cases}
\quad \Leftrightarrow \quad k(U,V) = f(v) + g(q) \ \forall\, V \in W
$$

is uniquely solvable and the solution fulfills the stability estimate

$$
\|U\|_W = \|\underline{u}\|_V + \|p\|_Q \leq c\left(\|f\|_{V'} + \|g\|_{Q'}\right)
$$

with the constant c depending only on α_a, α_b, β_a, β_b.

Our task now is to show stability, consistency and boundedness for the discrete problem to get uniquely solvability and the stability estimates. One essential ingredient for this is the treatment of the incompressibility constraint. The inf-sup condition on the discrete level is often the most important criteria for stability of numerical methods for incompressible flows. If α_b of Brezzi's Theorem or in our case $\alpha_{\mathcal{D}_h}$ is required to be independent of the mesh size the inf-sup criteria is most often called the LBB (Ladyshenskaja-Babuška-Brezzi) condition. We will again show consistency, stability, boundedness and approximation one after another before we conclude with a priori error estimates in a discrete broken H^1 norm.

2.3.2.1   Consistency

One important property of the discretization is the consistency, i.e. that the discrete equation is fulfilled if we replace the discrete solution with the exact one. The incompressibility condition is trivially fulfilled

as the exact solution is exactly divergence-free.

Let $(\underline{\mathbf{u}}, p) \in Z$ be the solution of (2.3.1). Then for all $q \in L^2(\Omega)$

$$\mathcal{D}_h(\underline{\mathbf{u}}, q) = \sum_{T \in \mathcal{T}_h} \int_T -\underbrace{\mathrm{div}(\underline{u})}_{=0} \; q \, d\underline{x} = 0 \tag{2.3.20}$$

And for $\mathcal{B}_h + \mathcal{D}_h$ we get for all $\underline{\mathbf{v}} \in S_0$:

$$
\begin{aligned}
&\mathcal{B}_h(\underline{\mathbf{u}}, \underline{\mathbf{v}}) + \mathcal{D}_h(\underline{\mathbf{v}}, p) \\
&= \sum_{T \in \mathcal{T}_h} \left\{ \int_T \nu \underline{\nabla}\, \underline{u} : \underline{\nabla}\, \underline{v} \, d\underline{x} - \int_{\partial T} \nu \frac{\partial \underline{u}}{\partial \underline{n}} \, [\![\underline{v}^t]\!] \, d\underline{s} - \int_{\partial T} \nu \frac{\partial \underline{v}}{\partial \underline{n}} \underbrace{[\![\underline{u}^t]\!]}_{=0} \, d\underline{s} \right. \\
&\qquad\qquad \left. + \int_{\partial T} \nu \tau_h \underbrace{[\![\underline{u}^t]\!]}_{=0} \, [\![\underline{v}^t]\!] \, d\underline{s} - \int_T \mathrm{div}(\underline{v}) \; p \, d\underline{x} \right\} \\
&= \sum_{T \in \mathcal{T}_h} \left\{ \int_T \mathrm{div}(-\nu \underline{\nabla}\, \underline{u}) \; \underline{v} \, d\underline{x} + \int_{\partial T} \nu \frac{\partial \underline{u}}{\partial \underline{n}} \underbrace{\left(\underline{v} - [\![\underline{v}^t]\!] \right)}_{=\underline{v}^n + \underline{v}_F^t} d\underline{s} \right. \\
&\qquad\qquad \left. + \int_T \underline{\nabla} p \; \underline{v} \, d\underline{x} \right\} - \sum_{E \in \mathcal{F}_h^{int}} \int_E \underbrace{[\![p]\!]^*}_{=0} \, \underline{v}_n \, d\underline{s} - \sum_{E \in \mathcal{F}_h^{ext}} \int_E p \; \underline{v}_n \, d\underline{s} \\
&\overset{(*)}{=} \sum_{T \in \mathcal{T}_h} \int_T \underbrace{\mathrm{div}(-\nu \underline{\nabla}\, \underline{u} + p \underline{I})}_{=\, \underline{f}} \; \underline{v} \, d\underline{x} + \sum_{E \in \mathcal{F}_h^{int}} \int_E \underbrace{[\![\nu \frac{\partial \underline{u}}{\partial \underline{n}}]\!]}_{=0} \left(\underline{v}^n + \underline{v}_F^t \right) d\underline{s} \\
&\qquad\qquad + \sum_{E \in \mathcal{F}_h^{ext}} \int_E \left(\nu \underline{\nabla}\, \underline{u} - p\, \underline{I} \right) \cdot \underline{n} \left(\underline{v}^n + \underline{v}_F^t \right) d\underline{s} \\
&= \int_\Omega \underline{f} \; \underline{v} \, d\underline{x} + \int_{\Gamma_D} \left(\nu \underline{\nabla}\, \underline{u} - p\, \underline{I} \right) \cdot \underline{n} \underbrace{\left(\underline{v}^n + \underline{v}_F^t \right)}_{=0} d\underline{s} \\
&\qquad\qquad + \int_{\Gamma_{out}} \underbrace{\left(\nu \underline{\nabla}\, \underline{u} - p\, \underline{I} \right) \cdot \underline{n}}_{=0} \left(\underline{v}^n + \underline{v}_F^t \right) d\underline{s} \\
&= \int_\Omega \underline{f} \; \underline{v} \, d\underline{x}
\end{aligned}
\tag{2.3.21}
$$

With those preparations we can conclude the following:

Proposition 2.3.3 (Galerkin Orthogonality). *Let $U_h = (\underline{\mathbf{u}}_h, p_h) \in Z_{h,D}$ be the solution of* (2.3.10) *and $U = (\underline{\mathbf{u}}, p) \in Z$ be the solution of* (2.3.1). *Then there holds*

$$\mathcal{B}_h(\underline{\mathbf{u}}_h - \underline{\mathbf{u}}, \underline{\mathbf{v}}_h) + \mathcal{D}_h(\underline{\mathbf{u}}_h - \underline{\mathbf{u}}, q_h) + \mathcal{D}_h(\underline{\mathbf{v}}_h, p_h - p) = 0 \quad \forall \, (\underline{\mathbf{v}}_h, q_h) \in S_{h,0} \times Q_h \tag{2.3.22a}$$

$$\Longleftrightarrow \qquad\qquad\qquad\qquad \mathcal{K}_h(U_h - U, V_h) = 0 \quad \forall \, V_h \in Z_{h,0} \tag{2.3.22b}$$

Proof. There holds $\mathcal{D}_h(\underline{\mathbf{u}}, q_h) = \mathcal{D}_h(\underline{\mathbf{u}}_h, q_h) = 0 \; \forall q_h \in Q_h$ (see eq. (2.3.20)). And with (2.3.21) we obtain

$$\mathcal{B}_h(\underline{\mathbf{u}}_h, \underline{\mathbf{v}}_h) + \mathcal{D}_h(\underline{\mathbf{v}}_h, p_h) - \mathcal{B}_h(\underline{\mathbf{u}}, \underline{\mathbf{v}}_h) - \mathcal{D}_h(\underline{\mathbf{v}}, p_h) = \int_\Omega \underline{f} \; \underline{v} \, d\underline{x} - \int_\Omega \underline{f} \; \underline{v} \, d\underline{x} = 0 \qquad \forall \, \underline{\mathbf{v}}_h \in S_{h,0} \tag{2.3.23}$$

which completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

2.3.2.2 Stability

The stability of the mixed problem consists of two parts, the (kernel-) coercivity of $\mathcal{B}_h$ and the *LBB* condition of $\mathcal{D}_h$. For $\mathcal{B}_h$ we can reuse the scalar results. For the *LBB* condition we will essentially benefit from the $H(\text{div})$-conformity and the interpolation properties presented in 2.2.5.
First, let us define discrete *HDG norms* analogously to the scalar case (see section 1.2.3.2):

$$\|\mathbf{u}\|_{1,*}^2 := \sum_{T \in \mathcal{T}_h} \left\{ \|\underline{\nabla}\, \underline{u}\|_T^2 + \frac{1}{h} [\![\mathbf{u}^t]\!]_{\partial T}^2 \right\} \tag{2.3.24}$$

$$\|\mathbf{u}\|_1^2 := \sum_{T \in \mathcal{T}_h} \left\{ \|\underline{\nabla}\, \underline{u}\|_T^2 + \frac{1}{h} [\![\mathbf{u}^t]\!]_{\partial T}^2 + h\|\frac{\partial \underline{u}}{\partial \underline{n}}\|_{\partial T}^2 \right\} \tag{2.3.25}$$

They are again equivalent on the discrete Space S_h which follows with the same line of argument as in the scalar case. In contrast to the scalar case we don't include a problem-dependent scaling like ε or ν which is owed to the different type of analysis we use in this chapter. Scalings with ν now appear explicitly in the estimates.

Proposition 2.3.4 (Coercivity). *For a shape regular mesh and $\tau_h h$ (with h the local mesh size) sufficiently large $\mathcal{B}_h(\cdot, \cdot)$ is coercive on S_h, that is*

$$\mathcal{B}_h(\mathbf{u}, \mathbf{u}) \geq c\nu \|\mathbf{u}\|_{1,*}^2 \geq \alpha_{\mathcal{B}_h} \nu \|\mathbf{u}\|_1^2 \qquad \forall\, \mathbf{u} \in S_h \tag{2.3.26}$$

with c, $\alpha_{\mathcal{B}_h} \in \mathbb{R}$ independent of the mesh size.

Proof. The proof can be straightforwardly taken from the proof of the scalar equivalent (Proposition 1.2.2). As the norms are accordingly defined all inequalities hold even with the same constants. Even the inverse inequality can be applied component-wise which leads to the same overall constant as in the scalar case. □

As the claim and the proof follow directly from the scalar case, even the stabilization parameter τ_h can be chosen in the same way as already discussed in chapter 1.

Proposition 2.3.5 (discrete LBB-condition for $\mathcal{D}_h$). *For $\mathcal{D}_h(\cdot, \cdot)$ there holds*

$$\sup_{\mathbf{u} \in S_h^{k+1}} \frac{\mathcal{D}_h(\mathbf{u}, q)}{\|\mathbf{u}\|_1} \geq \alpha_{\mathcal{D}_h} \|q\|_{L^2} \qquad \forall\, q \in Q_h^k \tag{2.3.27}$$

(as long as $\Gamma_D \neq \partial\Omega$ (see also the subsequent Remark 2.3.3)).

Proof. In short we use several relations to come to

$$\sup_{\mathbf{u} \in S_h^{k+1}} \frac{\mathcal{D}_h(\mathbf{u}, q)}{\|\mathbf{u}\|_1} \geq \frac{\mathcal{D}_h(\tilde{\mathbf{v}}, q)}{\|\tilde{\mathbf{v}}\|_1} \overset{2.}{\geq} \frac{\mathcal{D}_h(\mathbf{v}^*, q)}{c_2 \|\underline{v}^*\|_{H^1}} \overset{1.}{\geq} \frac{1}{c_1 c_2} \frac{\|q\|_{L^2}^2}{\|q\|_{L^2}} \geq \frac{1}{c_1 c_2} \|q\|_{L^2} \tag{2.3.28}$$

The idea of the inequality labeled with "1." is to construct an $[H^1]^d$-regular function $\underline{v}^*$ which fulfills $\text{div}(\underline{v}^*) = q$. In 2. we find a discrete function $\tilde{\mathbf{v}} \in S_h^{k+1}$ which fulfills $\mathcal{D}_h(\tilde{\mathbf{v}}, q) = \mathcal{D}_h(\mathbf{v}^*, q)$ and can be bounded in the *HDG norm* by the $H^1(\Omega)$ norm of $\underline{v}^*$. This procedure can also be found in [HL02]. We will explain the steps 1. and 2. in more detail now:

1. Choose $\underline{v}^*$, s.t. $\underline{v}^* = \underline{\nabla}\phi$ with $\phi \in H^2$ the solution of

$$-\Delta\phi = q \qquad \text{on } \Omega \quad \text{and} \qquad \frac{\partial\phi}{\partial\underline{n}} = 0 \quad \text{on } \partial\Omega \qquad (2.3.29)$$

The solution exists under appropriate regularity assumptions.
Then we use

(a) $\mathcal{D}_h(\underline{\mathbf{v}}^*, q) = \sum_{T \in \mathcal{T}_h} \int_T \underbrace{-\text{div}(\underline{v}^*)}_{-\Delta\phi = q} \, q \, d\underline{x} = \int_\Omega q^2 \, d\underline{x} = \|q\|_{L^2}^2$

(b) $\|\underline{v}^*\|_{H^1} \leq \|\phi\|_{H^2} \leq c_1\|q\|_{L^2}^2$ due to elliptic regularity

2. Now we use Lemma 2.3.1 which garantuees the existence of a discrete function $\tilde{\underline{v}} \in S_h^{k+1}$, s.t.

(a)
$$\mathcal{D}_h(\underline{\mathbf{v}}^*, q) = \mathcal{D}_h(\tilde{\mathbf{v}}, q) \; \forall \; q \in Q_h^k \qquad (2.3.30)$$

(b)
$$\|\tilde{\mathbf{v}}\|_1 \leq c_2\|\underline{v}^*\|_{H^1} \qquad (2.3.31)$$

$\square$

Remark 2.3.3 (Non-Uniqueness of the pressure if $\Gamma_D = \partial\Omega$):
Note, that if Dirichlet boundary conditions are prescribed on the whole boundary, $\mathcal{D}_h(\mathbf{v}, p) = 0$ for any constant pressure p. Thus the inf-sup condition does not work in this case. Then, the proof goes wrong as the Neumann problem (2.3.29) posed therein does not fulfill the compability condition $\int_\Omega q \, d\underline{x} = \int_{\partial\Omega} \frac{\partial\phi}{\partial\underline{n}} \, d\underline{s} = 0$ and thus has no solution.
The problem we observe here, is owed to the lack of uniqueness of the pressure of the Stokes problem. This problem obviously also exists on the continuous level (as p only appears as a $\underline{\nabla}p$ in the strong form).
We can solve this issue by forbidding constant-pressure contributions to p, i.e. replacing $L^2(\Omega)$ by $L_0^2(\Omega) := \{p \in L^2\Omega : \int_\Omega p \, d\underline{x} = 0\}$ in the weak formulation. Then, the pressure is made unique again. This also solves the problem with the inf-sup-condition, because if we replace Q_h^k by $Q_h^k \setminus \mathbb{R}$ in the discrete formulation, the compability condition for the Neumann problem (2.3.29) is fulfilled again and the proof works.

2.3.2.3   Boundedness

Proposition 2.3.6 (Boundedness). *For all $\mathbf{u}, \mathbf{v} \in S$ and $q \in Q$ there holds:*

$$|\mathcal{B}_h(\mathbf{u}, \mathbf{v})| \leq \beta_{\mathcal{B}_h} \, \nu \, \|\mathbf{u}\|_1 \, \|\mathbf{v}\|_1 \qquad (2.3.32)$$

with $\beta_{\mathcal{B}_h} = \sup_{\underline{x} \in \mathcal{F}_h}(1 + \tau_h h)$ and

$$|\mathcal{D}_h(\mathbf{u}, q)| \leq \underbrace{\sqrt{d}}_{= \beta_{\mathcal{D}_h}} \|q\|_{L^2} \, \|\mathbf{u}\|_1 \qquad (2.3.33)$$

Proof. (2.3.32) follows from the scalar poisson equation, where we used Proposition 1.2.3.
With $|\text{div}(\underline{u})| = |tr(\underline{\nabla} \, \underline{u})| \leq \sqrt{d}\|\underline{\nabla} \, \underline{u}\|_F$ and Cauchy Schwarz inequalities the second equation also holds true. $\square$

2.3.2.4 Approximation

Proposition 2.3.7 (Approximation of the divergence-free subspace). *Let $\underline{u} \in [H^1(\Omega)]^d \cap [H^m(\mathcal{T}_h)]^d$ be divergence-free, then the spaces S_h^* and S_h fulfill the following approximation relation*

$$\inf_{\underline{\mathbf{w}}_h \in S_h^*} \|\|\mathbf{w}_h - \mathbf{u}\|\|_1 \leq C \inf_{\underline{\mathbf{v}}_h \in S_h} \|\|\mathbf{v}_h - \mathbf{u}\|\|_1 \tag{2.3.34}$$

with C independent of the mesh size. This means that the approximation error *on the divergence-free subspace is (up to a constant) as small as the* approximation error *on the whole space.*

Proof. This proof is taken and adapted from [Brae97, remark 4.10]. As $\underline{\mathbf{v}}_h \in S_h$ there holds (see Lemma 2.3.1) $\underline{\Pi}_h^S(\underline{\mathbf{v}}_h) = \underline{\mathbf{v}}_h$ and for the solution $\underline{\mathbf{u}}$ there holds $\underline{\Pi}_h^S(\underline{\mathbf{u}}) \in S_h^*$. If we additionally make use of the boundedness of the projection $\underline{\Pi}_h^S$ we obtain:

$$\|\|\underline{\Pi}_h^S(\mathbf{u}) - \mathbf{u}\|\|_1 = \|\|\underline{\Pi}_h^S(\mathbf{u} - \mathbf{v}_h) - (\mathbf{u} - \mathbf{v}_h)\|\|_1 \leq \|\|\mathbf{u} - \mathbf{v}_h\|\|_1 + c\|\mathbf{u} - \mathbf{v}_h\|_{H(\mathsf{div})} \leq (1+c)\|\|\mathbf{u} - \mathbf{v}_h\|\|_1 \tag{2.3.35}$$

So for every arbitrary $\underline{\mathbf{v}}_h \in S_h$ we can construct a function $\underline{\mathbf{w}}_h = \underline{\Pi}_h^S(\underline{\mathbf{u}}) \in S_h^*$, s.t. the claim holds.

$\square$

Proposition 2.3.8. *On a shape regular mesh $\mathcal{T}_h$ which consists of affine transformed elements and a function $\underline{u} \in [H^1(\Omega)]^d \cap [H^m(\mathcal{T}_h)]^d$, $m = k+1 \geq 2$ there holds the following approximation result for S_h^k*

$$\inf_{\underline{\mathbf{v}}_h \in S_h^k} \|\|\mathbf{u} - \mathbf{v}_h\|\|_1^2 \leq C \sum_{T \in \mathcal{T}_h} h_T^{2k} \, |\underline{u}|_{H^m(T)}^2 \tag{2.3.36}$$

with C independent of the mesh size. For quasi-uniform meshes the direct conclusion of it is

$$\inf_{\underline{\mathbf{v}}_h \in S_h^k} \|\|\mathbf{u} - \mathbf{v}_h\|\|_1 \leq C \, h^k \, |\underline{u}|_{H^m(\mathcal{T}_h)} \tag{2.3.37}$$

Proof. Use Bramble-Hilbert Lemma with $L = \underline{\mathbf{id}}_h - \underline{\Pi}_{h,k}^S$ and a scalar product similar to the one used in the scalar case. Then, as in the scalar case we directly get the local estimate:

$$\|\|(\underline{\mathbf{id}}_h - \underline{\Pi}_{h,k}^S)\underline{u}\|\|_1^T \leq C \, |\underline{u}|_{H^m(T)} \qquad \forall \, \underline{u} \in [H^m(T)]^d \tag{2.3.38}$$

As the scaling of the norm is exactly the same as in the scalar case an adapted version of the proof of Proposition 1.2.6 also works in the vector-valued case. $\square$

As already pointed out we could replace $\Sigma_h^k = \mathsf{BDM}_k$ with $\Sigma_h^k = \mathsf{RT}_{k-1}$ to get a stable solution. But we would loose one order of accuracy in the last Proposition if we'd do so.

Proposition 2.3.9. *On a shape regular mesh $\mathcal{T}_h$ which consists of affine transformed elements and a function $p \in H^m(\mathcal{T}_h)$, $m = k+1 \geq 1$ there holds the following approximation result for Q_h^k as introduced in section 1.2.2.1.*

$$\inf_{q_h \in Q_h^k} \|p - q_h\|_{L^2}^2 \leq C \sum_{T \in \mathcal{T}_h} h_T^{2k+2} \, |p|_{H^m(T)}^2 \tag{2.3.39}$$

with C independent of the mesh size. For quasi-uniform meshes the direct conclusion of it is

$$\inf_{q_h \in Q_h^k} \|p - q_h\|_{L^2} \leq C \, h^{k+1} \, |p|_{H^m(\mathcal{T}_h)} \tag{2.3.40}$$

Proposition 2.3.10. *On a shape regular mesh $\mathcal{T}_h$ which consists of affine transformed elements and a function $(\underline{u}, p) \in [H^1(\mathcal{T}_h)]^d \cap [H^{m+1}(\mathcal{T}_h)]^d \times H^m(\mathcal{T}_h)$, $m = k + 1 \geq 1$ there holds the following approximation result*

$$\inf_{q_h \in Q_h^k} \|p - q_h\|_{L^2}^2 + \nu \inf_{\mathbf{v}_h \in S_h^{k+1}} \|\underline{\mathbf{u}} - \mathbf{v}_h\|_1^2 \leq Ch^{k+1} \left(|p|_{H^m(\mathcal{T}_h)}^2 + \nu \, |u|_{H^m(\mathcal{T}_h)}^2 \right) \tag{2.3.41}$$

2.3.2.5   Putting it all together

Lemma 2.3.11 (Cea's Lemma for Stokes). *Let $(\mathbf{u}, p) \in Z$ be the solution of* (2.3.1) *and* $(\underline{\mathbf{u}}_h, p_h) \in Z_h^{k+1}$ *the solution of* (2.3.10)*. Then there holds*

$$\|p - p_h\|_{L^2} + \sqrt{\nu}\|\underline{\mathbf{u}} - \mathbf{u}_h\|_1 \leq c \left\{ \inf_{q_h \in Q_h^k} \|p - q_h\|_{L^2} + \sqrt{\nu} \inf_{\mathbf{v}_h \in S_h^{k+1}} \|\underline{\mathbf{u}} - \mathbf{v}_h\|_1 \right\} \tag{2.3.42}$$

for c independent of the mesh size.

Proof. We basically use Cea's Lemma for $\mathcal{K}_h$ and again substitute $U := (\underline{\mathbf{u}}, p)$ and $V := (\mathbf{v}, q)$ and define $U_h, V_h, W_h \in Z_h^{k+1}$ accordingly . Due to Brezzi's Theorem, Lemma 2.3.2, we have inf-sup stability for $\mathcal{K}_h$ with a constant $\alpha_{\mathcal{K}_h}$. Boundedness is easily achieved: For all $U_h, V_h \in Z_h^{k+1}$:

$$
\begin{aligned}
|\mathcal{K}_h(U_h, V_h)| &\leq & |\mathcal{B}_h(\underline{\mathbf{u}}_h, \mathbf{v}_h)| &+& |\mathcal{D}_h(\underline{\mathbf{u}}_h, q_h)| &+& |\mathcal{D}_h(\mathbf{v}_h, p_h)| \\
&\leq & \beta_{\mathcal{B}_h} \nu \|\underline{\mathbf{u}}_h\|_1 \|\mathbf{v}_h\|_1 &+& \beta_{\mathcal{D}_h} \|p_h\|_{L^2} \|\mathbf{v}_h\|_1 &+& \beta_{\mathcal{D}_h} \|q_h\|_{L^2} \|\underline{\mathbf{u}}_h\|_1 \\
&\leq & \underbrace{\max(\beta_{\mathcal{B}_h}, \beta_{\mathcal{D}_h})}_{=\beta_{\mathcal{K}_h}} &\cdot& (\sqrt{\nu}\|\underline{\mathbf{u}}_h\|_1 + \|p_h\|_{L^2}) &\cdot& (\sqrt{\nu}\|\mathbf{v}_h\|_1 + \|q_h\|_{L^2}) \\
&= & \beta_{\mathcal{K}_h} \; \|U_h\|_{\mathcal{K}_h} \; \|V_h\|_{\mathcal{K}_h}
\end{aligned}
$$

$$\tag{2.3.43}$$

Now again we use the triangle inequality for the error $U - U_h$

$$\|U - U_h\|_{\mathcal{K}_h} \leq \|U - V_h\|_{\mathcal{K}_h} + \|V_h - U_h\|_{\mathcal{K}_h} \tag{2.3.44}$$

The second term can be bounded by the first term due to inf-sup stability, boundedness and consistency of $\mathcal{K}_h$. We take W_h such that inf-sup stability holds and V_h is an arbitrary element of Z_h^{k+1}:

$$
\begin{aligned}
\|U_h - V_h\|_{\mathcal{K}_h}\|W_h\|_{\mathcal{K}_h} &\leq & \frac{1}{\alpha_{\mathcal{K}_h}} \mathcal{K}_h(U_h - V_h, W_h) \\
&\leq & \frac{1}{\alpha_{\mathcal{K}_h}} \left(\underbrace{\mathcal{K}_h(U_h - U, W_h)}_{=0} + \mathcal{K}_h(U - V_h, W_h) \right) \\
&\leq & \frac{\beta_{\mathcal{K}_h}}{\alpha_{\mathcal{K}_h}} \|U - V_h\|_{\mathcal{K}_h}\|W_h\|_{\mathcal{K}_h}
\end{aligned}
\tag{2.3.45}
$$

So we have

$$\|U - U_h\|_{\mathcal{K}_h} \leq (1 + \frac{\beta_{\mathcal{K}_h}}{\alpha_{\mathcal{K}_h}})\|U - V_h\|_{\mathcal{K}_h} \tag{2.3.46}$$

which holds for all discrete functions V_h. So the limit is also true and if we keep in mind the definition of the $\|\cdot\|_{\mathcal{K}_h}$ norm the claim holds true with $c = (1 + \frac{\beta_{\mathcal{K}_h}}{\alpha_{\mathcal{K}_h}})$. Note that the constant scales in the same manner as for the scalar HDG discretization. $\square$

Lemma 2.3.12 (An estimate in the $\|\cdot\|_1$ norm)**.** *Let* $\mathcal{T}_h$ *be a quasi-uniform shape regular mesh,* $(\underline{\mathbf{u}}_h, p_h) \in Z_h^k$ *be the solution of* (2.3.10) *and* $(\underline{u}, p) \in [H^1(\Omega)]^d \cap [H^m(\mathcal{T}_h)]^d \times H^{m-1}(\mathcal{T}_h)$, $m \geq 2$ *the solution of* (2.3.1)*. Then there holds the following error estimate:*

$$\|p - p_h\|_{L^2} + \sqrt{\nu}\|\underline{\mathbf{u}} - \underline{\mathbf{u}}_h\|_1 \leq C h^s \left(\sqrt{\nu}|\underline{u}|_{H^s(\mathcal{T}_h)} + |p|_{H^{s-1}(\mathcal{T}_h)} \right) \qquad s = \min(k, m - 1) \quad (2.3.47)$$

Proof. Marriage Lemma 2.3.11 and the approximation results, i.e. Propositions 2.3.8 and 2.3.9.   □

Lemma 2.3.13 (Ceá-like Lemma for the velocity)**.** *Let* $(\underline{\mathbf{u}}, p) \in Z$ *be the solution of* (2.3.1) *and* $(\underline{\mathbf{u}}_h, p_h) \in Z_h$ *the solution of* (2.3.10)*. Then there holds the following estimate not including the pressure field:*

$$\|\underline{\mathbf{u}} - \underline{\mathbf{u}}_h\|_1 \leq c_1 \inf_{\underline{\mathbf{w}}_h \in S_h^*} \|\underline{\mathbf{u}} - \underline{\mathbf{w}}_h\|_1 \leq c_2 \inf_{\underline{\mathbf{w}}_h \in S_h} \|\underline{\mathbf{u}} - \underline{\mathbf{w}}_h\|_1 \tag{2.3.48}$$

with c_1 *and* c_2 *independent of the mesh size.*

Proof. Use coercivity, consistency $(\mathcal{B}_h(\underline{\mathbf{u}}_h - \underline{\mathbf{u}}, \underline{\mathbf{v}}_h) = -\mathcal{D}_h(\underline{\mathbf{v}}_h, p - p_h))$ and boundedness to get

$$\alpha_{\mathcal{B}_h} \, \nu \, \|\underbrace{\underline{\mathbf{u}}_h - \underline{\mathbf{w}}_h}_{=\underline{\mathbf{v}}_h}\|_1^2 \leq \mathcal{B}_h(\underline{\mathbf{u}}_h - \underline{\mathbf{w}}_h, \underline{\mathbf{v}}_h) = \underbrace{\mathcal{B}_h(\underline{\mathbf{u}}_h - \underline{\mathbf{u}}, \underline{\mathbf{v}}_h)}_{=-\mathcal{D}_h(\underline{\mathbf{v}}_h, p - p_h) = 0} + \mathcal{B}_h(\underline{\mathbf{u}} - \underline{\mathbf{w}}_h, \underline{\mathbf{v}}_h) \leq \beta_{\mathcal{B}_h} \nu \|\underline{\mathbf{u}} - \underline{\mathbf{w}}_h\|_1 \|\underline{\mathbf{v}}_h\|_1$$
$$(2.3.49)$$

where $\mathcal{D}_h(\underline{\mathbf{v}}_h, p - p_h) = 0$ just holds true on the (exactly) divergence-free subspace. Divide (2.3.49) by $\alpha_{\mathcal{B}_h}\|\underline{\mathbf{u}}_h - \underline{\mathbf{w}}_h\|_1$ and use the triangle inequality to get:

$$\|\underline{\mathbf{u}} - \underline{\mathbf{u}}_h\|_1 \leq \|\underline{\mathbf{u}} - \underline{\mathbf{w}}_h\|_1 + \|\underline{\mathbf{w}}_h - \underline{\mathbf{u}}_h\|_1 \leq \left(1 + \frac{\beta_{\mathcal{B}_h}}{\alpha_{\mathcal{B}_h}} \right) \|\underline{\mathbf{u}} - \underline{\mathbf{w}}_h\|_1 \tag{2.3.50}$$

which completes the proof.   □

With the last estimate we see that the approximation of the pressure field p is not at all required to fulfill any approximation condition explicitly to get optimal convergence results in the $\|\cdot\|_1$ norm for the velocity field $\underline{u}$. Nevertheless exactly divergence-free velocities are essential for the last estimate and so the pressure field has to be chosen rich enough. For the reduced basis presented before (see section 2.2.4.2), this means that only a constant pressure approximation (corresponding to the RT_0 DOF) works without destroying higher order convergence of the velocity field as long as the incompressibility constraint is ensured exactly (e.g. with the Finite Element space presented before).

We can show better results for the velocity field in the L^2 norm, if we use standard duality techniques. Therefore we have to assume regularity of the adjoint problem, which coincides with the primal problem except for the different force term $f \in L^2$. So, those assumptions are as realistic as in the estimates before.

Lemma 2.3.14 (Aubin Nitsche for the velocity)**.** *Let* $(\underline{\mathbf{u}}, p) \in Z$ *be the solution of* (2.3.1) *and* $(\underline{\mathbf{u}}_h, p_h) \in Z_h$ *the solution of* (2.3.10)*. Then there holds*

$$\|\underline{u} - \underline{u}_h\|_{L^2} \leq c \, h\|\underline{\mathbf{u}} - \underline{\mathbf{u}}_h\|_1 \tag{2.3.51}$$

for c *independent of the mesh size.*

Proof. We pose the adjoint stokes problem with $f(v) := (\underline{u} - \underline{u}_h, \underline{v})_\Omega$. Find $W \in Z$, s.t.

$$\mathcal{K}_h(V, W) = (\underline{u} - \underline{u}_h, \underline{v})_\Omega \qquad \forall\, V \in Z \tag{2.3.52}$$

Under regularity assumptions on the domain we may assume $W \in [H^2(\Omega)]^d \times H^1(\Omega)$. Then we choose $V = U - U_h = (\underline{u} - \underline{u}_h, p - p_h)$ and so get

$$\begin{aligned}
\|\underline{u} - \underline{u}_h\|_{L^2}^2 = \mathcal{K}_h(U - U_h, W) &\overset{(*)}{=} \mathcal{K}_h(U - U_h, W - W_h) \\
&= \mathcal{B}_h(\underline{u} - \underline{u}_h, \underline{w} - \underline{w}_h) + \underbrace{\mathcal{D}_h(\underline{w} - \underline{w}_h, p - p_h)}_{=0} + \underbrace{\mathcal{D}_h(\underline{u} - \underline{u}_h, q - q_h)}_{=0} \\
&\leq \beta_{\mathcal{B}_h} \nu \|\mathbf{u} - \mathbf{u}_h\|_1 \|\mathbf{w} - \mathbf{w}_h\|_1
\end{aligned} \tag{2.3.53}$$

where we shifted an arbitrary function $W_h \in Z^*$ into the bilinearform due to consistency in $(*)$. $\mathcal{D}_h(\cdot, \cdot) = 0$ holds as $\underline{u}$, $\underline{u}_h$, $\underline{w}$ and $\underline{w}_h$ are divergence-free. Now we can choose W_h as the linear projection of the velocity (which is divergence-free) and the constant projection of the pressure. As we have assumed sufficient regularity we have

$$\|\mathbf{w} - \mathbf{w}_h\|_1 \leq C_{appr}\, h\, |\underline{w}|_{H^2} \tag{2.3.54}$$

Furthermore the $|\cdot|_{H^2}$ semi-norm of $\underline{u}$ can be bounded by $\underline{u} - \underline{u}_h$ in the L^2 norm since we assumed sufficient regularity:

$$\nu |\underline{w}|_{H^2} \leq C_{reg} \|\underline{f}\|_{L^2} = C_{reg} \|\underline{u} - \underline{u}_h\|_{L^2} \tag{2.3.55}$$

Putting together (2.3.54) and (2.3.55) and dividing by $\|\underline{u} - \underline{u}_h\|_{L^2}$ the statement holds true with $c = \beta_{\mathcal{B}_h}\, C_{reg}\, C_{appr}$. □

We conclude the a priori error analysis section with the statement of optimal convergence rate for the velocity L^2 error:

Lemma 2.3.15 (L^2 norm estimate). *Let $\mathcal{T}_h$ be a quasi-uniform shape regular mesh, $(\mathbf{u}_h, p_h) \in Z_h^k$ be the solution of* (2.3.10) *and $(\underline{u}, p) \in [H^1(\Omega)]^d \cap [H^m(\mathcal{T}_h)]^d \times H^{m-1}(\mathcal{T}_h)$, $m \geq 2$ the solution of* (2.3.1). *Then there holds the following error estimate:*

$$\|\underline{u} - \underline{u}_h\|_{L^2} \leq C h^{s+1} \left(\sqrt{\nu} |\underline{u}|_{H^s(\mathcal{T}_h)} + |p|_{H^{s-1}(\mathcal{T}_h)} \right) \qquad s = \min(k, m-1) \tag{2.3.56}$$

Proof. Use Lemma 2.3.13, Lemma 2.3.14 and Proposition 2.3.8 □

2.4 Exactly divergence-free Hybrid DG method for the Oseen problem

In this section we want to add a vector-valued version of the scalar HDG upwind method to the Stokes Equations to include convective effects. We therefore start by replacing the nonlinear convective velocity $\underline{u}$ of the Navier-Stokes equations by a known velocity $\underline{w}$. Normally $\underline{w}$ is an old value of $\underline{u}$ during an iterative procedure. So the Oseen problem is

$$
\begin{cases}
\operatorname{div}(-\nu \underline{\nabla}\,\underline{u} + \underline{u} \otimes \underline{w} + p\underline{I}) &= \underline{f} & \text{in } \Omega \\
\operatorname{div}(\underline{u}) &= 0 & \text{in } \Omega \\
\underline{u} &= \underline{u}_D & \text{on } \Gamma_D \\
(\nu \underline{\nabla}\,\underline{u} - p\underline{I}) \cdot \underline{n} &= 0 & \text{on } \Gamma_{out}
\end{cases}
\tag{2.4.1}
$$

2.4.1 Introducing the method

Except for the (for now linear) convective term $\operatorname{div}(\underline{u} \otimes \underline{w})$ the numerical treatment of the set of equations is the same as for the Stokes Equations. We additionally introduce the bilinearform $\mathcal{C}_h(\underline{w}; \mathbf{u}, \mathbf{v})$ where $\underline{w}$ originates from a function $\mathbf{w} \in Z$. We use the same procedure to derive the Hybrid DG upwind bilinearform as in the scalar convective case by looking at each component one by one. The only difference appears when looking at the upwind-part. As $\underline{u}^n$ is continuous for each discrete function, we just have to treat the tangential component in an upwind fashion.

We directly give all three versions of the bilinearform as the derivation is straight-forward when knowing the scalar version and the notation used for the viscous bilinearform $\mathcal{B}_h$.

With $\quad \underline{u}^{up} = \underline{u}^n + \begin{cases} \underline{u}_F^t & w_n < 0 \\ \underline{u}^t & w_n > 0 \end{cases}$

$$
\begin{aligned}
&\mathcal{C}_h(\underline{w}, \mathbf{u}, \mathbf{v}) \\
&= \sum_{T \in \mathcal{T}_h} \left\{ -\int_T \underline{u} \otimes \underline{w} : \underline{\nabla}\,\underline{v}\, dx + \int_{\partial T} w_n \underline{u}^{up}\,\underline{v}\, ds + \int_{\partial T_{out}} w_n(\underline{u}_F^t - \underline{u}^t)\,\underline{v}_F^t\, ds \right\} \tag{2.4.2a} \\
&= \sum_{T \in \mathcal{T}_h} \left\{ \int_T \operatorname{div}(\underline{u} \otimes \underline{w})\,\underline{v}\, dx + \int_{\partial T_{in}} |w_n|(\underline{u}^t - \underline{u}_F^t)\,\underline{v}^t\, ds + \int_{\partial T_{out}} |w_n|(\underline{u}_F^t - \underline{u}^t)\,\underline{v}_F^t\, ds \right\} \tag{2.4.2b} \\
&= \sum_{T \in \mathcal{T}_h} \left\{ -\int_T \underline{u} \otimes \underline{w} : \underline{\nabla}\,\underline{v}\, dx + \int_{\partial T} w_n \underline{u}^{up}\, [\![\mathbf{v}^t]\!]\, ds + \int_{\partial T \cap \Gamma} w_n \underline{u}_F^t\,\underline{v}_F^t\, ds \right\} \tag{2.4.2c}
\end{aligned}
$$

So the discrete version of (2.4.1) is

$$
\text{Find } U = (\mathbf{u}, p) \in Z_{h,D},\, s.t. \quad \mathcal{K}_h(U, V) + \mathcal{C}_h(\underline{w}, \mathbf{u}, \mathbf{v}) = \langle \underline{f}, \underline{v} \rangle \quad \forall\, V = (\mathbf{v}, q) \in Z_{h,0} \tag{2.4.3}
$$

The bilinearform consisting of convection and viscosity(diffusion) is again called $\mathcal{A}_h : S \times S \to \mathbb{R}$:

$$
\mathcal{A}_h(\mathbf{u}, \mathbf{v}) := \mathcal{B}_h(\mathbf{u}, \mathbf{v}) + \mathcal{C}_h(\underline{w}, \mathbf{u}, \mathbf{v}) \tag{2.4.4}
$$

2.4.2 A priori error analysis

The analysis used in chapter 1 can, for the most part, be translated directly to the *Oseen problem* to show that $\mathcal{A}_h$ is consistent and coercive. Approximation results can be adopted from the last section

because we won't introduce another convection norm, but work with the discrete broken Sobolev norm $\|\cdot\|_1$. So the analysis here will be most meaningful in the diffusion dominant regions, i.e. where the Peclét number is small. The only ingredient we have to put some effort into is the boundedness of the bilinearform $\mathcal{C}_h$.

Consistency

The discretization of the vector-valued convection is also consistent:

Proposition 2.4.1 (Galerkin Orthogonality). *Let $(\underline{\mathbf{u}}_h, p_h) \in Z_{h,D}$ be the solution of (2.4.3) and $(\underline{\mathbf{u}}, p)$ be the solution of (2.4.1). Then there holds*

$$\mathcal{C}_h(\underline{w}; \underline{\mathbf{u}}_h - \underline{\mathbf{u}}, \underline{\mathbf{v}}_h) = 0 \qquad \forall \, \underline{\mathbf{v}}_h \in S_{h,0} \tag{2.4.5}$$

Proof. See the proof of the scalar analogous Proposition 1.2.1.   □

Stability

In this section we won't consider purely convective equations and so we always consider just the bilinearform $\mathcal{A}_h$ directly. That's why the used norms are taken from the diffusive limit (without viscous scaling), i.e. the discrete H^1-seminorm $\|\cdot\|_1$. To get stability, that is coercivity for this kind of analysis, is easy. From equation (1.3.18) we know that the contribution of the convective bilinearform is positive and so

$$\mathcal{A}_h(\underline{\mathbf{u}}, \underline{\mathbf{u}}) = \mathcal{B}_h(\underline{\mathbf{u}}, \underline{\mathbf{u}}) + \underbrace{\mathcal{C}_h(\underline{w}; \underline{\mathbf{u}}, \underline{\mathbf{u}})}_{\geq 0} \geq \alpha_{\mathcal{B}_h} \|\underline{\mathbf{u}}\|_1^2 \tag{2.4.6}$$

and we conclude:

Proposition 2.4.2 (Coercivity). *For a shape regular mesh and $\tau_h h$ (with h the local mesh size) sufficiently large $\mathcal{A}_h(\cdot, \cdot)$ is coercive on S_h, that is*

$$\mathcal{A}_h(\underline{\mathbf{u}}, \underline{\mathbf{u}}) \geq \alpha_{\mathcal{A}_h} \|\underline{\mathbf{u}}\|_1^2 \qquad \forall \, \underline{\mathbf{u}} \in S_h \tag{2.4.7}$$

with $\alpha_{\mathcal{A}_h} = \alpha_{\mathcal{B}_h} \nu \in \mathbb{R}$ only depending on the shape regularity of the mesh, the polynomial degree and the kinematic viscosity ν.

Boundedness

We now want the *trilinearform* $\mathcal{C}_h$ to be bounded (in all three arguments) w.r.t. the discrete norm $\|\cdot\|_1$. Unfortunately we can show this only in two dimensions:

Proposition 2.4.3. *For $\underline{\mathbf{u}}, \underline{\mathbf{w}} \in (\Sigma \cap [H^1(\mathcal{T}_h)]^2) \times F^t$, $div(\underline{w}) = 0$ and $\underline{\mathbf{v}} \in S_h$ there holds*

$$|\mathcal{C}_h(\underline{w}, \underline{\mathbf{u}}, \underline{\mathbf{v}})| \leq \beta_{\mathcal{C}_h}^* \|\underline{\mathbf{w}}\|_1 \|\underline{\mathbf{u}}\|_1 \|\underline{\mathbf{v}}\|_1 = \beta_{\mathcal{C}_h} \|\underline{\mathbf{u}}\|_1 \|\underline{\mathbf{v}}\|_1 \tag{2.4.8}$$

for $\beta_{\mathcal{C}_h}^* \in \mathbb{R}$ and $\beta_{\mathcal{C}_h} = \beta_{\mathcal{C}_h}^* \|\underline{\mathbf{w}}\|_1$ independent of the mesh size.

Proof. See appendix, B   □

Remark 2.4.1:
The assumption $\underline{\mathbf{u}}, \underline{\mathbf{w}} \in (\Sigma \cap [H^1(\mathcal{T}_h)]^d) \times F^t$ is satisfied for discrete functions of S_h as well as for all functions of $[H^1(\Omega)]^d$ if we again use the tangential trace operator to identify a function of $[H^1(\Omega)]^d$ with one of F^t. Thus the discrete and the continuous solution can be plugged in here for $\underline{\mathbf{u}}$ and $\underline{\mathbf{w}}$. Actually for $\underline{\mathbf{w}}$ the tangential component is not important at all, s.t. the regularity and the used norm for $\underline{\mathbf{w}}$ could be optimized further and a wider class of functions (e.g. functions without facet-values) could be used. We indicate that by not using the bold writing for the first argument of $\mathcal{C}_h$. However, in our case $\underline{\mathbf{w}}$ is going to be an approximation to $\underline{\mathbf{u}}$, s.t. using the same regularity assumptions and norms is reasonable.

Putting it all together
As we now still have coercivity, boundedness and consistency with respect to the "old" norm $\|\|\cdot\|\|_1$, we can use the results we got for the Stokes problem. Thus optimal convergence in the broken H^1 norm as well as L^2 norm for the velocity field are achieved. Nevertheless there remains a disadvantage of the analysis we used here:
If we look at the proof of Lemma 2.3.11 or Lemma 2.3.13 we have to replace the constants for $\mathcal{B}_h$ with the constants for $\mathcal{A}_h$. But this results in

$$\left(1 + \frac{\beta_{\mathcal{A}_h}}{\alpha_{\mathcal{A}_h}}\right) = \left(1 + \frac{\nu\beta_{\mathcal{B}_h} + \|\|\mathbf{w}\|\|_1 \beta^*_{\mathcal{C}_h}}{\nu\alpha_{\mathcal{B}_h}}\right) = c_1 + c_2 \cdot \frac{\|\|\mathbf{w}\|\|_1}{\nu} \qquad \text{for } c_1, c_2 \in \mathbb{R} \tag{2.4.9}$$

Thus for convection dominating over viscosity (i.e. diffusion) the constant increases. That's why this kind of analysis, which is based upon ellipticity of the bilinearform $\mathcal{A}_h$ is less appropriate for large convection. Nevertheless existence and uniqueness of the discrete solution is still ensured in that case! Actually the method itself works fine also in the convection dominated case. As long as viscosity dominates the behaviour of the system the constants which appear have the same magnitude as in the Stokes case and the results are comparable also w.r.t. the constant.

2.5   Exactly divergence-free Hybrid DG method for the Navier-Stokes problem

If we replace the convective velocity of the Oseen problem we recover the original steady Navier-Stokes problem:

$$\begin{cases} \operatorname{div}(-\nu\underline{\nabla}\,\underline{u} + u \otimes u + p\underline{I}) & = \underline{f} & \text{in } \Omega \\ \operatorname{div}(\underline{u}) & = 0 & \text{in } \Omega \\ \underline{u} & = \underline{u}_D & \text{on } \Gamma_D \\ (\nu\underline{\nabla}\,\underline{u} - p\underline{I}) \cdot \underline{n} & = 0 & \text{on } \Gamma_{out} \end{cases} \quad (2.5.1)$$

This set of nonlinear equation can then be solved iteratively by means of a Picard iteration, which solves an Oseen problem with the convective velocity of the last iteration at each step. Here the solution of the Stokes problem is a good initial guess as it garantues a divergence-free velocity.
The discrete problem reads

$$\text{Find } U = (\mathbf{u}, p) \in Z_{h,D}, s.t. \quad \mathcal{K}_h(U,V) + \mathcal{C}_h(\underline{u}, \mathbf{u}, \mathbf{v}) = \langle \underline{f}, \underline{v} \rangle \quad \forall\, V = (\mathbf{v}, q) \in Z_{h,0} \quad (2.5.2)$$

2.5.1   Introducing the method

The only step we have to take to get from the discretization of the Oseen problem to the nonlinear problem is to describe an iterative procedure converging to the discrete solution. E.g. this iterative procedure could be a Newton-Raphson method. Here, we consider the simpler approach, the Picard-iteration:

Algorithm 2.5.1 Picard-iteration for the steady Navier-Stokes equations

---
Set $i = 0$
Solve the discrete steady Stokes problem (2.3.10) for $\mathbf{u}^0$
repeat
   increase i by 1
   Solve the discrete steady Oseen problem (2.4.3) for $\mathbf{u}^n$ with $\mathbf{w} = \mathbf{u}^{n-1}$
until desired accuray is achieved ($\|\underline{u}^n - \underline{u}^{n-1}\|_{L^2} \leq \epsilon$)

Now we hope that this fixpoint iteration converges. If it does, it obviously converges to a discrete solution of (2.5.2). Convergence of the fixpoint iteration, which would imply existence and uniqueness of (2.5.2), is only ensured under so called "smallness assumptions", i.e. volume forces and boundary conditions are assumed to be sufficiently small. To use those smallness assumptions is reasonable as existence and uniqueness of the continuous problem also depends on them. In [CKS05, Theorem 4.7] and [CKS05, Theorem 4.8] those smallness assumptions where used to proof existence and uniqueness of the discrete solution for a similar exactly divergence-free DG method. They also provide a priori error estimates which ensure the optimal convergence properties we achieved for the Stokes problem and the Oseen problem. The same technique could be applied here. A detailed look at those proofs shows that except for adapting the bilinearforms and norms used therein the proof can actually be directly applied to our discretization.

2.5.2 Numerical Example: Kovasznay flow

We consider an exact solution to the Navier-Stokes equations obtained by Kovasznay [Kova48] which we will use as a reference solution. The Kovasznay flow is two-dimensional and describes the fluid flow behind a grid. The solution is given by

$$\underline{u}(x,y) = \begin{pmatrix} 1 - e^{\lambda x}\cos(2\pi y) \\ \frac{\lambda}{2\pi}e^{\lambda x}\sin(2\pi y) \end{pmatrix}$$

$$p(x,y) = -\frac{1}{2}e^{2\lambda x} + \overline{p} \qquad \text{with } \overline{p} \in \mathbb{R}$$

with

$$\lambda = \frac{-8\pi^2}{\nu^{-1} + \sqrt{\nu^{-2} + 64\pi^2}}$$

On the boundary of the domain $\Omega = [-\frac{1}{2}, \frac{3}{2}] \times [0,2]$ we prescribe inhomogeneuos Dirichlet data given by the exact solution. If we additionally demand the average of p $\left(\frac{1}{|\Omega|} \int_\Omega p\, dx\right)$ to be zero the solution of the steady Navier-Stokes equations is unique and is exactly the solution proposed by Kovasznay with a suitable $\overline{p}$.

Again we choose τ_h as in the numerical examples before with α to be 2 and consider the case where $\nu = 1$. The chosen meshes we used and a view on the solution can be found in Figure 2.5.1.

(a) Initial and finest mesh used for the discretization and absolute value of the velocity solution

(b) Absolute value of the velocity solution and streamlines
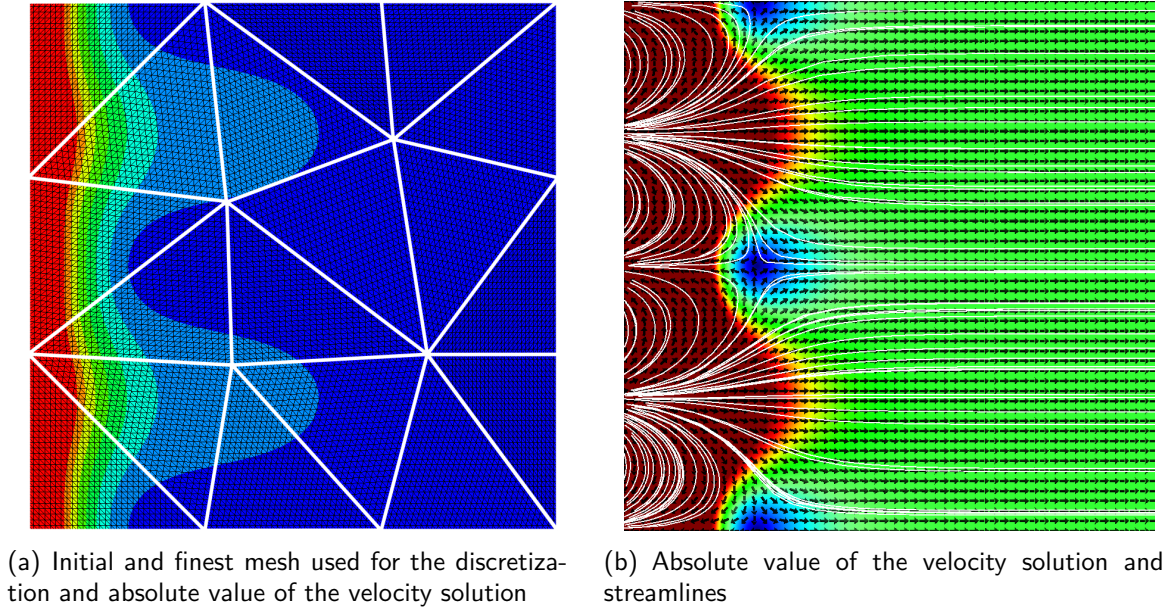
Figure 2.5.1: Solution and meshes for the Kovasznay problem

Convergence of the method

In the next table we see the convergence of the method for the chosen example. We observe that the pressure error in the L^2-norm and velocity error $\underline{e}_h$ in the broken H^1 seminorn converge with the expected rate h^p whereas the L^2-norm of the velocity error $\underline{e}_h$ converges even an order faster. That is what we expected from the analysis carried out before. We can furthermore observe how the reduction of the Finite Element space reflects on the number of unknowns (compare DOF with DOF*) and the

number of nonzero elements in the matrices (compare nze, which stands for "nonzero entries", with nze*). For $p = 2$ the unknowns are reduced roughly by a factor of $\frac{1}{4}$ and a factor of nearly $\frac{1}{2}$ is achieved for $p = 4$. We see that the asymptotial reduction factor of $\frac{1}{2}$ is already achieved for moderately high p. For all meshes and polynomial degrees the iteration counts are bounded by 10.

meshlevel (elements)	p	DOF	DOF*	$\|\underline{\mathbf{e}}_h\|_{L^2}$ error	order	$\|\mathrm{div}(\mathbf{u}_h)\|_{L^2}$ error	$\|\underline{\nabla}(\underline{\mathbf{e}}_h)\|_{L^2}$ error	order	$\|p - p_h\|_{L^2}$ error	order
0 (18)	2	306	234	3.21	—	6.77e-9	5.76e1	—	4.19e1	—
1 (72)		1 152	864	6.25e-1	2.36	5.48e-9	1.91e1	1.60	2.16e1	0.96
2 (288)		4 464	3 312	8.62e-2	2.86	5.20e-9	5.22	1.87	6.52	1.73
3 (1 152)		17 568	12 960	1.00e-2	3.11	5.69e-9	1.31	1.99	1.75	1.90
4 (4 608)		69 696	51 264	1.17e-3	3.10	7.83e-9	3.26e-1	2.01	4.49e-1	1.96
5 (18 432)		277 632	203 904	1.42e-4	3.04	1.28e-8	8.11e-2	2.01	1.13e-1	1.99

Table 2.2: Errors and order of convergence of pressure and velocity in different norms as well as a comparison between DOF of the hybrid divergence-free DG method with and without the reduced basis for the kovasznay flow problem using polynomial spaces up to degree 2.

meshlevel (elements)	p	DOF	DOF*	$\|\underline{\mathbf{e}}_h\|_{L^2}$ error	order	$\|\mathrm{div}(\mathbf{u}_h)\|_{L^2}$ error	$\|\underline{\nabla}(\underline{\mathbf{e}}_h)\|_{L^2}$ error	order	$\|p - p_h\|_{L^2}$ error	order
0 (18)	4	780	456	3.24e-1	—	2.70e-9	9.36	—	3.8e1	—
1 (72)		3 000	1 704	1.56e-2	4.38	4.86e-9	8.59e-1	3.45	1.22	4.96
2 (288)		11 760	6 576	5.62e-4	4.79	5.03e-9	6.06e-2	3.83	1.09e-1	3.48
3 (1 152)		46 560	25 824	1.84e-5	4.93	5.80e-9	3.89e-3	3.96	7.83e-3	3.80
4 (4 608)		185 280	102 336	5.73e-7	5.01	1.88e-9	2.43e-4	4.00	5.08e-4	3.95
5 (18 432)		739 200	407 424	1.80e-8	4.99	7.81e-8	1.52e-5	4.00	3.19e-5	3.99

Table 2.3: Errors and order of convergence of pressure and velocity in different norms as well as a comparison between DOF of the hybrid divergence-free DG method with and without the reduced basis for the kovasznay flow problem using polynomial spaces up to degree 4.

Chapter 3

Time Integration for incompressible flow problems

3.1 Semi-discrete form of convection diffusion type problems

So far we only considered steady flow problems in the last chapters in order to concentrate on the proposed spatial discretization. Now we to take time discretization into consideration. We will use the method of lines here, i.e. we use spatial discretization and apply time discretization schemes afterwards. As we discussed convection diffusion and the Navier-Stokes Equations we want to cast both problems into a unified form. The discretization of the unsteady problems (1.4.1) and (2.5.1) then read

$$M\left(\frac{\partial u}{\partial t}, v\right) + B(u,v) + C(u;u,v) = f(v) \qquad \forall \, v \in V \tag{3.1.1}$$

or in operator notation:

$$M\frac{\partial u}{\partial t} + Bu + C(u)u = f \tag{3.1.2}$$

In the following table we summarize the interpretations of this notation which result either in the unsteady convection diffusion equation or the Navier-Stokes equations.

unified notation	scalar convection diffusion counterpart	Navier-Stokes counterpart
$u, v \in V$	scalar quantity $\mathbf{u} = (u, u_F) \in V_h$	velocity-pressure pair $U{=}(\underline{u}, \underline{u}_F, p){\in}Z_h$
$M(\cdot, \cdot)$	Mass matrix $M(u,v) := \displaystyle\int_\Omega u\, v\, dx$	Mass matrix $M(u,v) := \displaystyle\int_\Omega \underline{u}\, \underline{v}\, dx$
matrix properties	block diagonal	not block diagonal
$B(\cdot, \cdot)$	diffusion bilinearform $\mathcal{B}_h(\cdot, \cdot)$ (1.2.8)	stokes bilinearform, i.e. viscosity, pressure and incompressibility constraint: $\mathcal{K}_h(\cdot, \cdot)$ (2.3.11)
matrix properties	symmetric positive definite	symmetric indefinite
$C(\cdot; \cdot, \cdot)$	convection bilinearform $\mathcal{C}_h(\cdot, \cdot)$ (1.3.4)	nonlinear vector-valued convection bilinearform $\mathcal{C}_h(\cdot; \cdot, \cdot)$ (2.4.2)
properties	linear	nonlinear
$f(\cdot)$	source terms	force terms

We see that the bilinearforms for the Navier-Stokes equations are less comfortable as those of the convection diffusion equation. Actually due to the incompressibility constraint the system of equations arising from spatial discretization is not even an ordinary system of equations, but a differential algebraic (DAE) system. That is why we will concentrate on this more involved case, transferring the results to convection diffusion is then straight forward.

Remark 3.1.1 (On facet unknowns):
Time derivatives of facet unknowns do not appear in either of both formulations and thus even the convection diffusion discretization is not an ordinary system of equations (the equations for v_F are algebraic). But as facet unknowns can be eliminated easily, see section 1.2.2.2, 1.4.1.1 and 1.5.1, this problem is only a formal one and can be overcomed easily. In the proceeding we will neglect to adress this issue.

3.2   Approaches for Time Integration: explicit vs. implicit

We now want to discuss the numerical integration of (3.1.2). Therefore let's have a look on the two most famous lowest order time integration methods, the explicit (forward) and the implicit (backward) Euler method. Denoting with u^n, u^{n+1} the values of the discrete solutions at time t^n and t^{n+1}, respectively, and furthermore defining the increment $\Delta u := u^{n+1} - u^n$, the explicit Euler method reads:

$$M\frac{\Delta u}{\Delta t} = f^n - Bu^n - C(u^n)u^n \tag{3.2.1}$$

whereas the implicit Euler method is

$$M\frac{\Delta u}{\Delta t} = f^{n+1} - Bu^{n+1} - C(u^{n+1})u^{n+1} \tag{3.2.2}$$

Those methods are the simplest representatives for explicit and implicit time integration methods whose advantages and disadvantages will be discussed in the following.

3.2.1   Explicit methods

For the Navier-Stokes equations fully explicit methods are not directly possible as the incompressibility constraint is algebraic and thus has to be treated implicitely. At least you need to treat the pressure and the incompressibility constraint implicitely and the rest explicitely.
A more elegant way to derive explicit methods for Navier-Stokes equations is the projection onto the divergence-free subspace. On the divergence-free subspace, the Navier-Stokes equations are a system of ordinary differential equations. For this we can apply every explicit method directly. Afterwards we have to go back on the full space by reintroducing the incompressibility constraint for the new time step, i.e. an implicit term. As this is nevertheless the minimalistic approach for time integration of Navier-Stokes equations we may still call it explicit.
Explicit time integration comes with basically two nice properties: Discrete systems arising from explicit time integration are linear and normally involve only the mass matrix. For Navier-Stokes equations

the situation is - as we have seen - more involved and the mass matrix has to be extended with the incompressibility constraint:

$$\begin{pmatrix} M & 0 \\ 0 & 0 \end{pmatrix} \frac{\Delta u}{\Delta t} = \dots \qquad \longrightarrow \qquad \begin{pmatrix} M & D^T \\ D & 0 \end{pmatrix} \frac{\Delta u}{\Delta t} = \dots$$

where M is the mass matrix and D^T, D the discrete gradient of p and the discrete incompressibility constraint, respectively.

Remark 3.2.1 (Modifications of the spatial discretization for explicit time integration):
Solving linear systems of equations with the extended mass matrix

$$\begin{pmatrix} M & D^T \\ D & 0 \end{pmatrix}$$

for the discretization proposed in chapter 2 is not at all cheap (mainly due to the fact that M^{-1} is not sparse), at least we can do better. Therefore let's shortly explain possible modifications of the formulation which make the use of the spatial discretization more efficient for explicit time integration: If we use completely discontinuous finite elements instead of $H(\text{div})$-conforming ones and introduce lagrange multipliers p_F, q_F to ensure normal-continuity, the mass matrix M is block diagonal and the schur complement $DM^{-1}D^T$ is a sparse symmetric positive definite matrix which is comparable to a discrete laplacian for the pressure. As $H(\text{div})$-conformity is then ensured through the additional lagrange multipliers the spatial discretization coincides with the one proposed in chapter 2.

The major disadvantage of explicit methods is that they are only conditionally stable, s.t. a certain time step restriction has to be fulfilled. Due to the time step restriction this approach renders inefficient for non-negligibly large viscosity. Those time step restrictions behave in the following way:
Convection owes time step restrictions which scale linearly with the spatial resolution and the magnitude of the convective term $|u|$

$$\Delta t \leq c_C \frac{h}{p^2} \frac{1}{|u|} \tag{3.2.3}$$

for some constant $c_C \in \mathbb{R}$ whereas the viscosity demands more:

$$\Delta t \leq c_B \frac{h^2}{p^4} \frac{1}{\nu} \tag{3.2.4}$$

for some constant $c_B \in \mathbb{R}$. The latter condition is often the more disturbing one as it scales quadratically with $\frac{h}{p^2}$ and dominates over the first condition for moderate values of ν.

3.2.2 Implicit methods

In contrast to explicit methods most practically relevant implicit methods are unconditionally stable. This of course is desireable, but it doesn't come for free. The system of equation you have to solve in each time step may be considerably expensive: In our case it is nonlinear and each linearization results in a linear system of equations with a nonsymmetric and indefinite matrix. This implies that the matrices are, in contrast to explicit methods, varying at each time step.
For stiff systems implicit methods are indispensable and the viscosity matrix of the Navier-Stokes discretization is considerably stiff for moderate values of ν which results in the inconvenient time step restriction (3.2.4) for explicit methods. Thus for problems with non-small diffusity (viscosity) implicit methods should be prefered.

3.2.3   Semi-Implicit methods

Let's also discuss a mixture which combine the features of explicit and implicit methods with the following aim:

- Avoid solving nonlinear, nonsymmetric equations (too expensive)

- Avoid quadratical dependence of the time step on the mesh size (see (3.2.4))

or, to put the second point in another way:

- Avoid integration of the stiff part with an explicit scheme

We achieve this by splitting the r.h.s. of (3.1.2) into an implicit and an explicit part. The diffusive term, i.e. B is treated implicitely and the (nonlinear) convective term C is treated explicitely resulting in the first order *semi-implicit Euler*:

$$M\frac{\Delta u}{\Delta t} = f^n - Bu^{n+1} - C(u^n)u^n$$
$$\Leftrightarrow (M + \Delta t B)\Delta u = \Delta t(f^n - Bu^n - C(u^n)u^n)$$

(3.2.5)

As f is (assumed to be) independent of u the choice of the time level for f is arbitrary, i.e. the costs for evaluation are the same for t^n and t^{n+1}. In the following we will treat f explicitely, i.e. in the same way as the convective term. Nevertheless other choices are possible.

3.2.4   Stability vs. Efficiency

In practice all three approaches discussed above are use. Which of those methods is the most appropriate, i.e. the more efficient, normally depends on the spatial discretization, the flow parameters and perhaps also on the computer you are working on[1]. Implicit methods are most robust and could be used for all practically relevant flow configurations, but if the use of explicit or semi-implicit methods is possible without dramatical restrictions of the time step, the implicit method is often (far) more expensive. In general there can't be stated a best choice. Nevertheless as (fully) explicit and (fully) implicit approaches are well-known we will concentrate on the discussion of semi-implicit approaches.

3.3   Higher Order methods

All three time discretizations mentioned above ((3.2.1), (3.2.2) and (3.2.5)) are only of first order. For fully explicit (and with this respect the method with implicit treatment of the incompressibility still applies as fully explicit) and fully implicit schemes extensions to higher order methods by e.g. Runge-Kutta formulas are well-known. For semi-implicit schemes the extension to higher order seems to be less famous. We want to discuss the so called implicit-explicit (IMEX) schemes which allow for such an extension in the next section.

---

[1]e.g. some methods may be easier to parallalize then other

3.4 IMEX Time Integration

Starting point is the equation (3.1.2)

$$M\frac{\partial u}{\partial t} = f - Bu - C(u)u$$

where B is a linear operator describing a diffusive process and thereby is stiff, whereas $C(u)$ is a possibly nonlinear convective operator. In the following we want to treat the diffusive operator always implicitly and the convective operator always explicitly. For convenience we thus write (remember that f is also treated explicitly)

$$M\frac{\partial u}{\partial t} + Bu = f - C(u)u$$

Terms on the right hand side are treated explicitly and those on the left hand side implicitly. We already saw the first order version of these IMEX schemes. It is obtained by replacing the time derivative by a finite difference and by exchanging u by u^{n+1} on the l.h.s. and u by u^n on the r.h.s. resulting in

$$M\frac{\Delta u}{\Delta t} + Bu^{n+1} = f^n - C(u^n)u^n$$

We will discuss two different realizations of higher order extensions in the sequel. These are either achieved with special *partitioned Runge-Kutta methods* or *Multistep methods*. Further we will focus on methods which are *stiffly accurate* as they garantuee that the solution at the new time level is divergence-free. DAE systems can be seen, with this regard, as a limit case of stiff systems and so *stiffly accurate* schemes are appropriate for those kind of problems.

3.4.1 IMEX Runge-Kutta methods

The basic idea of IMEX Runge-Kutta methods is simple: combine two *compatible* Runge-Kutta schemes from which one is implicit and one is explicit, in a way such that a scheme is achieved which has the overall order of accuracy which coincides with the minimal order of accuracy of the individual Runge-Kutta schemes.

To find the compatibility condition for two Runge-Kutta schemes to match in the desired sense is actually quite simple: They just have to work on the same time levels. Methods fulfilling these criteria were introduced and discussed in [ARS97].

In the following we will only consider the pair of Runge-Kutta schemes which consists of one explicit Runge-Kutta scheme and one *diagonally implicit* Runge-Kutta scheme which is *stiffly accurate*. Furthermore we prefer methods which are *singly diagonally implicit* and L-stable[2]. A method is called *diagonally implicit* if the according butcher tableau has only zero entries above the diagonal, i.e. each stage value does not depend on future stage values. It is furthermore called *singly diagonally implicit* if the diagonal entries of the butcher tableau are all the same. If we have *singly diagonally implicit* methods and the stiff part, i.e. the part we want to treat implicitly, is linear, the linear equations we have to solve within one step use the same matrix. This allows for reusing factorizations and preconditioners. If the time step is additionally used for multiple time steps the l.h.s. matrix can be used even more often. We require *stiffly accurate* schemes to deal with the incompressibility-constraint. Each new time step has to fulfill it exactly and not only in an interpolated sense. *Stiffly accurate* schemes

---

[2]Thus all examples we present here are singly diagonally implicit and L-stable what does not hold in general for all methods fitting into the presented framework.

can always be written, s.t. the last stage value of the Runge-Kutta scheme can be taken as the new instance for the new time step t^{n+1}.

The described IMEX Runge-Kutta methods (diagonally implicit & stiffly accurate) work as showed in algorithm 3.4.1.

Algorithm 3.4.1 Stiffly accurate s stage IMEX Runge-Kutta Time Integration
---
1: Use Startvalue u^n
2: Evaluate $C^1 = C(u^n)u^n - f(t^n)$
3: **for** $i = 1$ to s do :
4:     Solve $(M + \Delta t \overleftarrow{\mathbf{a}}_{\mathbf{i,i}} B)u^i = Mu^n - \Delta t \sum_{j=1}^{i} \overrightarrow{\mathbf{a}}_{\mathbf{i+1,j}} C^j - \Delta t \sum_{j=1}^{i-1} \overleftarrow{\mathbf{a}}_{\mathbf{i,j}} B^j$ for u^i
5:     Evaluate $B^i = Bu^i$
6:     Evaluate $C^{i+1} = C(u^i)u^i - f(t^n + \mathbf{c_i}\Delta t)$
7: **end for**
8: Set $u^{n+1} = u^s$

Coefficients with left-sided arrows refer to a Butcher tableau of a (diagonally) implicit Runge-Kutta scheme and those with a right-sided arrow are taken from a Butcher tableau of a compatible explicit s-stage Runge-Kutta scheme.

Normally Butcher tableaus also include coefficients b_j which describe which linear combination of stage values should be used for the new instance u^{n+1}. Here, the last stage value is automatically the new value at time level t^{n+1}, s.t. the coefficients $\overrightarrow{\mathbf{a}}_{\mathbf{s,j}}$ and $\overleftarrow{\mathbf{a}}_{\mathbf{s,j}}$ coincide with $\overrightarrow{\mathbf{b}}_{\mathbf{j}}$ and $\overleftarrow{\mathbf{b}}_{\mathbf{j}}$, respectively. In table 3.1 the general pair of Butcher tableaus for an s-stage IMEX Runge-Kutta schemes is shown.

	explicit						implicit				
0	0	0	0	$\dots$	0	0					
$\mathbf{c_1}$	$\overrightarrow{\mathbf{a}}_{\mathbf{2,1}}$	0	0	$\dots$	0	0	$\overleftarrow{\mathbf{a}}_{\mathbf{1,1}}$	0	$\dots$	0	0
$\mathbf{c_2}$	$\overrightarrow{\mathbf{a}}_{\mathbf{3,1}}$	$\overrightarrow{\mathbf{a}}_{\mathbf{3,2}}$	0	$\dots$	0	0	$\overleftarrow{\mathbf{a}}_{\mathbf{2,1}}$	$\overleftarrow{\mathbf{a}}_{\mathbf{2,2}}$	$\ddots$	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\ddots$	$\vdots$
$\mathbf{c_{s-1}}$	$\overrightarrow{\mathbf{a}}_{\mathbf{s,1}}$	$\overrightarrow{\mathbf{a}}_{\mathbf{s,2}}$	$\overrightarrow{\mathbf{a}}_{\mathbf{s,3}}$	$\dots$	0	0	$\overleftarrow{\mathbf{a}}_{\mathbf{s-1,1}}$	$\overleftarrow{\mathbf{a}}_{\mathbf{s-1,2}}$	$\dots$	$\overleftarrow{\mathbf{a}}_{\mathbf{s-1,s-1}}$	0
$\mathbf{c_s}$	$\overrightarrow{\mathbf{a}}_{\mathbf{s+1,1}}$	$\overrightarrow{\mathbf{a}}_{\mathbf{s+1,2}}$	$\overrightarrow{\mathbf{a}}_{\mathbf{s+1,3}}$	$\dots$	$\overrightarrow{\mathbf{a}}_{\mathbf{s+1,s}}$	0	$\overleftarrow{\mathbf{a}}_{\mathbf{s,1}}$	$\overleftarrow{\mathbf{a}}_{\mathbf{s,2}}$	$\dots$	$\overleftarrow{\mathbf{a}}_{\mathbf{s,s-1}}$	$\overleftarrow{\mathbf{a}}_{\mathbf{s,s}}$
	$\|$	$\|$	$\|$	$\dots$	$\|$	$\|$	$\|$	$\|$	$\dots$	$\|$	$\|$
	$\overrightarrow{\mathbf{b}}_{\mathbf{1}}$	$\overrightarrow{\mathbf{b}}_{\mathbf{2}}$	$\overrightarrow{\mathbf{b}}_{\mathbf{3}}$	$\dots$	$\overrightarrow{\mathbf{b}}_{\mathbf{s}}$	$\overrightarrow{\mathbf{b}}_{\mathbf{s+1}}$	$\overleftarrow{\mathbf{b}}_{\mathbf{1}}$	$\overleftarrow{\mathbf{b}}_{\mathbf{2}}$	$\dots$	$\overleftarrow{\mathbf{b}}_{\mathbf{s-1}}$	$\overleftarrow{\mathbf{b}}_{\mathbf{s}}$

Table 3.1: IMEX Runge-Kutta Butcher tableau consisting of an embedded pair of Butcher tableaus for compatible Runge-Kutta methods

The coefficients $\mathbf{c_i}$ coincide for both single Butcher tableau, s.t. both Runge-Kutta methods work on the same time levels, thus both methods are *compatible*. We see on the left side the Butcher tableau of the explicit Runge-Kutta method embedded in the IMEX Runge-Kutta Butcher tableau. The coefficients of the original scheme are extended by $\overrightarrow{\mathbf{a}}_{\mathbf{s+1,j}} = \overrightarrow{\mathbf{b}}_{\mathbf{j}}$. The embedding of the implicit Runge-Kutta method is straight forward. Now the new Butcher tableau reflects the procedure of algorithm 3.4.1:

In the first row we have only zeros, i.e. only the convective part at t^n is evaluated. Then the next row

describes which linear combination of convection terms and diffusion terms (including the new stage value) should be used to calculate the new stage value. Continuing until stage s is reached gives us an approximation for u at t^{n+1}. Therefore no additional linear combination of old stage values has to be computed as $\overleftarrow{\mathbf{a}}_{\mathbf{s,j}} = \overleftarrow{\mathbf{b}}_{\mathbf{j}}$ due to the stiffly accuracy of the implicit method.

Note also that lines 5 and 6 of algorithm 3.4.1 don't have to be executed for $i = s$.

If the operator B is zero we recover the explicit Runge-Kutta method and for $C = 0$ we recover the (diagonally) implicit Runge-Kutta method.

Let's discuss the lowest order case $s = 1$. Here, we have exactly four free parameters, these are $\mathbf{c_1}$, $\overrightarrow{\mathbf{a}}_{\mathbf{1,1}}$, $\overleftarrow{\mathbf{a}}_{\mathbf{1,1}}$ and $\mathbf{b_1}$. For consistency (of the explicit method) we need $\mathbf{b_1} = \overrightarrow{\mathbf{a}}_{\mathbf{1,1}} = 1$. This corresponds to the use of the explicit euler for the explicit part. For our problems (f is not treated implicitely and no other term depends directly on t) the choice of $\mathbf{c_s}$ is arbitrary (see algorithm 3.4.1). So the only important parameter we have is to choose is $\overleftarrow{\mathbf{a}}_{\mathbf{1,1}} =: \theta$. For $\theta \in (0, 1]$ we recover the implicit θ-scheme for the implicit part. For $\theta \geq \frac{1}{2}$ the implicit method is unconditionally stable and at least of first order (for $\theta = \frac{1}{2}$ we recover the second order accurate Crank-Nicholson method). So both methods are of first order and compatible. For $\theta = 1$ we recover the semi-implicit Euler method which is completely described with the following Butcher tableau:

$$
\begin{array}{c}
\text{Implicit Euler} \quad \rightarrow \quad \text{Semi-Implicit Euler} \quad \leftarrow \quad \text{Explicit Euler}
\end{array}
$$

$$
\begin{array}{c|c}
1 & 1 \\ \hline
 & 1
\end{array}
\qquad \rightarrow \qquad
\begin{array}{c|cc|c}
0 & 0 & 0 & \\
1 & 1 & 0 & 1 \\ \hline
 & 1 & 0 & 1
\end{array}
\qquad \leftarrow \qquad
\begin{array}{c|c}
0 & 0 \\ \hline
 & 1
\end{array}
$$

Table 3.2: The semi-implicit Euler as a 1-stage IMEX Runge-Kutta scheme

We want to complete the presentation of IMEX Runge-Kutta methods by stating second and third order accurate IMEX Runge-Kutta methods which use SDIRK (singly diagonally implicit Runge-Kutta) schemes. Both are taken from [ARS97].

For second order accuracy it is sufficient to use a 2-stage IMEX Runge-Kutta scheme with the following butcher tableau:

$$
\begin{array}{c}
\text{((SA)SDIRK2):} \quad \rightarrow \qquad\qquad \text{(SA)RKIMEX2} \qquad\qquad \leftarrow \qquad\qquad \text{ERK2}
\end{array}
$$

$$
\begin{array}{c|cc}
\gamma & \gamma & 0 \\
1 & 1-\gamma & \gamma \\ \hline
 & 1-\gamma & \gamma
\end{array}
\;\rightarrow\;
\begin{array}{c|ccc|cc}
0 & 0 & 0 & 0 & & \\
\gamma & \gamma & 0 & 0 & \gamma & 0 \\
1 & \delta & 1-\delta & 0 & 1-\gamma & \gamma \\ \hline
 & \delta & 1-\delta & 0 & 1-\gamma & \gamma
\end{array}
\;\leftarrow\;
\begin{array}{c|cc}
0 & 0 & 0 \\
\gamma & \gamma & 0 \\ \hline
 & \delta & 1-\delta
\end{array}
$$

Table 3.3: 2-stage second order accurate IMEX Runge-Kutta scheme

To achieve third order we have to use at least $s = 4$ stages. This is due to the combination of consistency and compatibility conditions that have to be fulfilled. One possibility of a 4-stage third order accurate IMEX Runge-Kutta scheme is given next.

0	0	0	0	0	0				
$1/2$	$1/2$	0	0	0	0	$1/2$	0	0	0
$2/3$	$11/18$	$1/18$	0	0	0	$1/6$	$1/2$	0	0
$1/2$	$5/6$	$-5/6$	$1/2$	0	0	$1/2$	$1/2$	$1/2$	0
1	$1/4$	$7/4$	$3/4$	$-7/4$	0	$3/2$	$-3/2$	$1/2$	$1/2$
	$1/4$	$7/4$	$3/4$	$-7/4$	0	$3/2$	$-3/2$	$1/2$	$1/2$

Table 3.4: 4-stage third order accurate Runge-Kutta IMEX scheme

In [KC03] IMEX Runge-Kutta methods with 6 stages which has overall accuracy order 4 is given as well as an alternative for the third order accurate IMEX Runge-Kutta method. Both use an ESDIRK method which in contrast to the SDIRK methods we stated has one explicit first stage even for the implicit method. This reduces the computational effort one step further[3]. Furthermore those methods use embedded Runge-Kutta formulas which allow for error estimation and adaptation as well as dense output. For readers with solid background in numerical methods for ordinary differential equations [KC03] also provides deeper insight into generalizations of IMEX schemes (to splitting of the operator into more than two parts) and order and coupling conditions and their derivations. In [KCGH07] these methods are used to solve compressible flow problems.

3.4.2   Multistep IMEX methods

Another possibility to generalize the first order semi-implicit Euler scheme to higher order accuracy is the use of multistep methods. The basic idea is to extrapolate information from old time steps to achieve higher order accuracy. Thus the number of function evaluations and linear systems which have to be solved is reduced to one for both, which makes the implementation efficient and straight forward: Extrapolated values for the stiff, the nonstiff and the mass term replace the evaluations at time level t^n of the semi-implicit Euler method and with the coefficient c^* denoting the weight of the stiff term at time level t^{n+1} we obtain:

$$(M + \Delta t c^* B)u^{n+1} = M \sum_{j=0}^{s-1} a_j u^{n-j} + \Delta t \sum_{j=0}^{s-1} b_j C(u^{n-j}) + \Delta t \sum_{j=0}^{s-1} c_j B(u^{n-j})$$

The following table shows possible choices up to order 4 which are taken from [ARW95]. Notice that SBDF methods can be generated generically of arbitrary order by left-sided finite differences, but as the BDF method is only L-stable up to order 4, the SBDF method is less desirable for $p > 4$.

---

[3]Thus the alternative third order IMEX Runge-Kutta method only has to solve three systems of equations

Name	s	c^*	a_0	a_1	a_2	a_3	b_0	b_1	b_2	b_3	c_0	c_1	c_2	c_3
TFE(θ)	1	θ	1				1				$1-\theta$			
BEFE, [1],SBDF1	1	1	1				1				0			
CNFE[1/2]	1	1/2	1				1				1/2			
[[γ,c]]	2	$\frac{\gamma+c/2}{\gamma+1/2}$	$\frac{2\gamma}{\gamma+1/2}$	$\frac{-\gamma+1/2}{\gamma+1/2}$			$\frac{\gamma+1}{\gamma+1/2}$	$-\frac{\gamma}{\gamma+1/2}$			$\frac{1-\gamma-c}{\gamma+1/2}$	$\frac{c}{2\gamma+1}$		
CNAB [[1/2,0]]	2	1/2	1	0			3/2	$-1/2$			1/2	0		
MCNAB [[1/2,1/8]]	2	9/16	1	0			3/2	$-1/2$			3/8	1/16		
CNLF [[0,1]]	2	1	0	1			2	0			0	1		
SBDF2 [[1,0]]	2	2/3	8/3	$-2/3$			2	-1			0	0		
SBDF3	3	6/11	18/11	$-9/11$	2/11		18/11	$-18/11$	6/11		0	0	0	
SBDF4	4	12/25	48/25	$-36/25$	16/25	$-3/25$	48/25	$-72/25$	48/25	$-12/25$	0	0	0	0

Table 3.5: Some Multistep IMEX schemes (source: [ARW95])

Some explanations to the table: The first order 1-stage IMEX multistep methods can be parametrized by one parameter θ. The first method is the "Theta-Forward-Euler" (TFE) method for general θ. The "Semi-implicit Backward Differentiation Formula" of first order (SDBF1) which is exactly the "Backward-Euler-Forward-Euler" (BEFE) is then again the "Theta-Forward-Euler" method for $\theta = 1$, whereas the "Crank-Nicholson-Forward-Euler"(CNFE) is the TFE method for $\theta = 1/2$.

The set of all second order 2-stage IMEX multistep methods can be parametrized by 2 parameters γ and c. Again, the general case ist presented first and then four special cases, namely the "Crank-Nicholson-Adams-Bashforth" (CNAB), the "Modified-Crank-Nicholson-Adams-Bashforth" (MCNAB), the "Crank-Nicholson-Leap-Frog" (CNLF) and the "Semi-implicit Backward Differentiation Formula" of second order (SBDF2) follow. As mentioned before the SBDF method can be derived for arbitrary order, but only the case $s \le 4$ is relevant for practical applications, as L-stability is lost for larger s.

Although multistep methods are obviously computationally cheap (at the cost of a bit more memory to keep old time values), you can not, in contrast to the Runge-Kutta methods presented before, expect a gain in the time step restriction. On the contrary: the time step restriction gets worse for higher order multistep methods. So the apparent efficiency is compensated. In our calculations IMEX Runge-Kutta methods have been more robust and also a bit cheaper than IMEX Multistep methods.

Appendix A

Notation, elementary inequalities and orthogonal polynomials

A.1   Notation

In this section the notation used through out the diploma thesis is summarized.

General		Abbreviations	
characteristic element length, e.g. diameter	: h		
vector	: $\underline{v}$		
outer normal	: $\underline{n}$	Degree(s) of freedom	: DOF
compound function	: $\mathbf{v} = (v, v_F)$	Finite Element Method	: FEM
compound vector function	: $\underline{\mathbf{v}} = (\underline{v}, \underline{v}_F^t)$	Continuous Galerkin	: CG
		Discontinuous Galerkin	: DG
velocity pressure pair	: $U = (\underline{u}, \underline{u}_F^t, p)$	Hybrid Disc. Galerkin	: HDG
dyadic product	: $\otimes \quad (\underline{a} \otimes \underline{b})_{i,j} = a_i b_j$	Hybrid mixed	: HM
smaller up to const.	: $\preceq$	Cauchy Schwarz (ineq.)	: C.S.
greater up to const.	: $\succeq$		
smaller and greater up to constant	: $\simeq$		

Constants		Triangulation	
stability constant for bilinearform $\mathcal{A}_h$	: $\alpha_{\mathcal{A}_h}$	Triangulation	: $\mathcal{T}_h$
boundedness constant for bilinearform $\mathcal{A}_h$	: $\beta_{\mathcal{A}_h}$	set of all element boundaries	: $\mathcal{F}_h$
shape regularity constant on one element	: c_T	set of interior el. interfaces	: $\mathcal{F}_h^{int}$
shape regularity for the whole mesh	: c_{sr}	set of exterior el. interfaces	: $\mathcal{F}_h^{ext}$

Norms and scalar products

For $f, g \in L^2(\Omega)$ we use the following notation for the L^2 inner product and the L^2 norm:

$$\begin{aligned}
(f,g)_\Omega &:= \int_\Omega f(x)\, g(x)\, d\underline{x} \\
\|f\|_\Omega^2 &:= \int_\Omega f(x)^2\, d\underline{x}
\end{aligned}$$

DG and HDG operators

As we consider having double and even triple-valued functions on interfaces, we have to distinguish between values which come from functions of different sides and functions which live only on the interface. The interface functions are always indicated by the index F. When we work with interface unknowns, neighbouring unknowns don't interact directly, so that for this case the following jump operator which is single-sided and an according norm is sufficient:

$$\begin{aligned}
[\![\mathbf{f}]\!] &:= f - f_F \\
[\![\mathbf{f}^t]\!] &:= \underline{f}^t - \underline{f}_F^t \\
[\![f]\!]_{\partial\Omega} &:= \|f - f_F\|_{\partial\Omega}
\end{aligned}$$

If we are interested in average values and jump operators between neighbouring values without incorperating interface unknowns [1], we define for a scalar function f and a vector function $\underline{h}$ - with $\underline{n}^+$, $\underline{n}^-$ the outwards pointing normals of each side - the following average and jump operators:

$$\begin{aligned}
\{\!\{f\}\!\} &:= \frac{1}{2}(f^+ + f^-) \\
[\![f]\!]^* &:= f^+ - f^- \\
\{\!\{\underline{h}\}\!\}^* &:= \underline{h}^+\, \underline{n}^+ - \underline{h}^-\, \underline{n}^- \\
[\![\underline{h}]\!]_n &:= \underline{h}^+\, \underline{n}^+ + \underline{h}^-\, \underline{n}^-
\end{aligned}$$

All those operators are scalar values. The stars indicate that the signs of those values are direction-dependent[2].

---

[1] like it is the case for *Standard DG*

[2] depending on which side is marked with $+$ and which with $-$

Infinite dimensional spaces	**Finite Element spaces**
	$\mathcal{P}^k(T) = \left\{ \sum\limits_{i=0}^{k} \sum\limits_{\substack{l,m,n \geq 0 \\ l+m+n=k}} a_{l,m,n} x^l y^m z^n, a_{l,m,n} \in \mathbb{R} \right\}$
$Q = L^2(\Omega)$	$Q_h^k = \left\{ u : u \in \mathcal{P}^k(T) \; \forall \; T \in \mathcal{T}_h \right\}$
$F = L^2(\mathcal{F}_h)$	$F_h^k = \left\{ u_F : u_F \in \mathcal{P}^k(E) \; \forall \; E \in \mathcal{F}_h \right\}$
$F^t = \left\{ \underline{u}_F^t : \underline{u}_F^t \in [L^2(\mathcal{F}_h)]^d \times \underline{n} \right\}$	$F_h^{k,t} = \left\{ \underline{u}_F : \underline{u}_F \in [\mathcal{P}^k(E)]^d \times \underline{n} \; \forall \; E \in \mathcal{F}_h \right\}$
$V = \left\{ (u, u_F) : u \in H^2(\mathcal{T}_h) \cap H^1(\Omega), \right.$	$V_h^k = Q_h^k \times F_h^k$
$\qquad \left. u_F \in L^2(\mathcal{F}_h) \right\}$	
$V_D = \left\{ (u, u_F) \in V, u_F = u_D \text{ on } \Gamma_D \right\}$	$V_{h,D} = \left\{ (u, u_F) \in V_h, u_F = u_D \text{ on } \Gamma_D \right\}$
$V_0 = \left\{ (u, u_F) \in V, u_F = 0 \text{ on } \Gamma_D \right\}$	$V_{h,0} = \left\{ (u, u_F) \in V_h, u_F = 0 \text{ on } \Gamma_D \right\}$
$\Sigma = H(\mathsf{div}, \Omega)$	$\Sigma_h^k = \left\{ \underline{u} : \underline{u} \in [\mathcal{P}^k(T)]^d \; \forall \; T \in \mathcal{T}_h, \right.$
	$\qquad \left. [\![\underline{u}]\!]_n = 0 \; \forall \; E \in \mathcal{F}_h \right\} = \mathsf{BDM}_k$
$S = \Sigma \times F^t$	$S_h^k = \Sigma_h^k \times F_h^{k,t}$
$Z = [H^1(\Omega)]^d \cap [H^2(\mathcal{T}_h)]^d \times L^2(\Omega)$	$Z_h^k = S_h^k \times Q_h^{k-1}$
$W = H^1(\Omega)$	$W_h^k = \left\{ u; u \in \mathcal{P}^k(T) \; \forall \; T \in \mathcal{T}_h, [\![u]\!]^* = 0 \; \forall \; E \in \mathcal{F}_h \right\}$

Projections

L^2 projection on all facets of one element ∂T : $\Pi_{h,k}^{fac}$

L^2 projection on one element T : $\Pi_{h,k}^{el}$

L^2 projection on V_h : $\mathbf{\Pi}_{h,k}$

$H(\mathsf{div})$-conforming projection on Σ_h : $\underline{\underline{\Pi}}_{h,k}^{\Sigma}$

L^2 projection on F_h^t : $\Pi_{h,k}^{f,t}$

$H(\mathsf{div})$-conforming projection on S_h : $\underline{\underline{\Pi}}_{h,k}^{S}$

A.2 Elementary Inequalities

Cauchy Schwarz inequality

Let V be a vector space over $\mathbb{R}$ with the inner product $(a, b)_N$ for $a, b \in V$. Then

$$(a, b)_N \leq \|a\|_N \|b\|_N$$

holds with $\|x\|_N = \sqrt{(x, x)_N}$. This inequality is useful for different norms on different vector spaces. In many proofs (especially boundedness proofs) we make use of two special cases of it:

a) In the first version we use the Cauchy Schwarz inequality for the L^2 scalar product on Ω

$$(f, g)_\Omega := \int_\Omega f(x) g(x) \, d\underline{x}$$

Thus for all $f, g \in V = L^2(\Omega)$ there holds:

$$\int_\Omega f(x) g(x) \, d\underline{x} = (f, g)_\Omega \leq \|f\|_\Omega \|g\|_\Omega = \sqrt{\int_\Omega f^2(x) \, d\underline{x}} \sqrt{\int_\Omega g^2(x) \, d\underline{x}} \qquad \text{(A.2.1a)}$$

b) The second version appears often when sums (over 3 or 4 components or over the triangulation) are involved. The l^2 scalar product is defined as

$$(a, b)_2 = \sum_{i=1}^{m} a_i \, b_i$$

for $a, b \in \mathbb{R}^m$. Thus for sums there holds

$$\sum_{i=1}^{m} a_i \, b_i = (a, b)_2 \leq \|a\|_2 \|b\|_2 = \sqrt{\sum_{i=1}^{m} a_i^2} \sqrt{\sum_{i=1}^{m} b_i^2} \qquad \text{(A.2.1b)}$$

Young's Inequality

Let $a, b, \gamma \in \mathbb{R}$. Then there holds

$$0 \leq (a - \gamma b)^2 = a^2 - 2a\gamma b + (\gamma b)^2$$

s.t. with straight forward manipulations we get Young's inequality:

$$\Rightarrow ab \leq \frac{a^2}{2\gamma} + \frac{\gamma b^2}{2} \qquad \text{(A.2.2)}$$

Inverse Inequalities

At several places we will make use of the inverse inequality which bounds the L^2 norm on the boundary by the L^2 norm at the domain for polynomials. Let u be a polynomial in $\underline{x} \in \mathbb{R}^d, d = 1, 2, 3$ with order p on a polyhedra T. Then there holds

$$\|u\|_{\partial T} \leq \frac{c}{\sqrt{h}} \|u\|_T \tag{A.2.3a}$$

with $c \in \mathbb{R}$ independent of the diameter of the polyhedra T.

Especially for the normal derivatives of polynomials we will use this inequality. That's why we also give the following direct conclusion:

$$\left\| \frac{\partial u}{\partial \underline{n}} \right\|_{\partial T} \leq \|\underline{\nabla} u\|_{\partial T} \leq \frac{c}{\sqrt{h}} \|\underline{\nabla} u\|_T \tag{A.2.3b}$$

The constant c does not depend on the mesh size but it clearly depends on p, as the variation from boundary surfaces to the inner volume may increase with rising polynomial degree. For clarity let's also cite an expression for $\frac{c}{\sqrt{h}}$ on a simplex from [WH03]. The equation (A.2.3a) holds true for

$$\frac{c}{\sqrt{h}} = \sqrt{\frac{(p+1)(p+d)}{d}} \sqrt{\frac{|\partial T|}{|T|}}$$

A.3   Orthogonal polynomials

Legrendre polynomials denoted by l_i are orthogonal polynomials spanning $\mathcal{P}^p([-1, 1])$. They satisfy

$$(l_i, l_j)_{L^2([-1,1])} = \frac{2}{2i+1} \delta_{i,j}$$

with $\delta_{i,j}$ the kronecker symbol. They can be computed with the follow recurrence relations:

$$\begin{aligned} l_0(x) &= 1, \\ l_1(x) &= x, \\ (n+1)l_{n+1}(x) &= (2n+1)l_n(x)x - nl_{n-1}(x), \qquad n \geq 1 \end{aligned}$$

Integrated Legrendre polynomials denoted by L_i are defined as follows:

$$L_n(x) := \int_{-1}^{x} l_{n-1}(\xi) \, d\xi \qquad \text{for } x \in [-1, 1] \text{ and } n \geq 2$$

They are mutually orthogonal with respect to the H^1-seminorm, i.e.

$$\int_{-1}^{1} L_i'(x) L_j'(x) \, dx = 0 \qquad \text{for } i \neq j$$

The polynomials L_n, $n \geq 2$ vanish at the interval bounds, i.e. $L_n(-1) = L_n(1) = 0$ and span $\mathcal{P}_0^p([-1, 1])$, that is the space of all polynomials up to degree p which have zero values on the bounds. For those polynomials there also holds a three-term-recurrence relation:

$$\begin{aligned} L_1(x) &= x, \\ L_2(x) &= \frac{1}{2}(x^2 - 1), \\ (n+1)L_{n+1}(x) &= (2n-1)L_n(x)x - (n-2)L_{n-1}(x), \qquad n \geq 2 \end{aligned}$$

Note that L_1 was added to allow for a general recursion relation although there does hold $L_1(x) \neq \int_{-1}^{x} l_0(\xi)\, d\xi$ An interesting property of the integrated Legendre polynomials is that they are almost orthogonal also with respect to the L^2 inner product:

$$(L_i, L_j)_{L^2([-1,1])} = 0 \qquad \text{for } |i - j| > 2$$

Scaled Legrendre polynomials denoted by $l_i^{\mathcal{S}}$ are defined as follows:

$$l_n^{\mathcal{S}}(x, t) := t^n l_n\left(\frac{x}{t}\right) \qquad \text{for } x \in [-1, 1],\ t \in (0, 1]$$

Thanks to the multiplication with the monomial t^n, the functions are free of fractions and stay polynomial of order n, i.e. $l_n^{\mathcal{S}} \in \mathcal{P}^n([-1, 1] \times [0, 1])$ and the limit $t \to 0$ is well defined. We again get three-term recurrences:

$$
\begin{aligned}
l_0^{\mathcal{S}}(x, t) &= 1, \\
l_1^{\mathcal{S}}(x, t) &= x, \\
(n + 1)l_{n-1}^{\mathcal{S}}(x, t) &= (2n + 1)l_n^{\mathcal{S}}(x, t)x - nt^2 l_{n-1}^{\mathcal{S}}(x, t), \qquad n \geq 2
\end{aligned}
$$

Scaled Integrated Legrendre polynomials denoted by $L_i^{\mathcal{S}}$ are defined as

$$
\begin{aligned}
L_n^{\mathcal{S}}(x, t) &:= t^n L_n\left(\frac{x}{t}\right) \qquad \text{for } x \in [-1, 1], t \in (0, 1] \\
&= \int_{-t}^{x} l_n^{\mathcal{S}}(x, t)\, d\xi \qquad \text{for } x \in [-1, 1], t \in (0, 1] \text{ and } n \geq 2
\end{aligned}
$$

and they satisfy $L_n^{\mathcal{S}} \in \mathcal{P}_0^n([-1, 1] \times [0, 1])$ and the following three-term-recurrence relations:

$$
\begin{aligned}
L_1(x)^{\mathcal{S}} &= x, \\
L_2(x^{\mathcal{S}}) &= \frac{1}{2}(x^2 - t^2), \\
(n + 1)L_{n+1}^{\mathcal{S}}(x) &= (2n - 1)L_n^{\mathcal{S}}(x)x - (n - 2)t^2 L_{n-1}^{\mathcal{S}}(x), \qquad n \geq 2
\end{aligned}
$$

Appendix B

Selected Proofs

Proof of Lemma 2.3.1

Proof. The proof is similar to the LBB-proof used in [HL02], but has some modifications to account for the additional facet functions.

Let's define $\underline{\mathbf{w}} := \underline{\mathbf{\Pi}}_k^S \underline{u}$. Now start with the definition of the norm and work on each term one after another.

$$\|\underline{\mathbf{\Pi}}_k^S \underline{u}\|_1 = \|\underline{\mathbf{w}}\|_1 = \sum_{T \in \mathcal{T}_h} \|\nabla \underline{w}\|_T^2 + \frac{1}{h} [\![\underline{\mathbf{w}}^t]\!]_{\partial T}^2 + h \| \frac{\partial \underline{w}}{\partial \underline{n}} \|_{\partial T}^2 = ...$$

As $\underline{w}$ is piecewise polynomial we can use inverse estimates to absorb the last term by the first one

$$... \leq \sum_{T \in \mathcal{T}_h} c_1 \|\nabla \underline{w}\|_T^2 + \frac{1}{h} [\![\underline{\mathbf{w}}^t]\!]_{\partial T}^2 = ...$$

For the first term we can use boundedness results of the $H(\text{div})$-conforming interpolation. The second part is more tricky. To separate the element functions and the facet functions we shift the function $\underline{u}$ in by using the triangle inequality

$$... \leq \sum_{T \in \mathcal{T}_h} c_1 \|\nabla \underline{w}\|_T^2 + \frac{1}{h} \|\underline{w}^t - \underline{u}^t\|_{\partial T}^2 + \frac{1}{h} \|\underline{u}^t - \underline{w}_F^t\|_{\partial T}^2 = ...$$

For the last part we can use the L^2-projection approximation property ($k \geq 1$)

$$\|\underline{u}^t - \underline{w}_F^t\|_{\partial T}^2 \leq c_2 h^2 \|\underline{u}\|_{\partial T}^2$$

and with the trace inequality [BS94, Theorem 1.6.6.] we additionally get

$$\|\underline{u}\|_{\partial T}^2 \leq \frac{c_3}{h} \|\underline{u}\|_T \|\underline{u}\|_{H^1(T)} \leq \frac{c_3}{h} \|\underline{u}\|_{H^1(T)}^2$$

This leaves us with ($c_4 = c_2 c_3$)

$$... \leq \sum_{T \in \mathcal{T}_h} c_1 \|\nabla \underline{w}\|_T^2 + c_4 \|\underline{u}\|_{H^1(T)}^2 + \frac{1}{h} \|\underline{w}^t - \underline{u}^t\|_{\partial T}^2 = ...$$

We now concentrate on the last problem, i.e. to bound $\|\underline{w}^t - \underline{u}^t\|_{\partial T}^2$. Let's therefore substitute $\underline{v} := \underline{w}^t - \underline{u}^t$. Now we use the same trace inequality that we used before, but transform to the reference element $\hat{T}$ before and back afterwards:

$$\|\underline{v}\|_{\partial T}^2 \leq c_5 h^{d-1} \|\underline{v}\|_{\partial \hat{T}}^2 \leq c_6 \left(h^{\frac{d}{2}} \|\underline{v}\|_{\hat{T}} \right) \left(h^{\frac{d-2}{2}} \|\underline{v}\|_{H^1(\hat{T})} \right) \leq c_9 \|\underline{v}\|_T (\frac{1}{h} \|\underline{v}\|_T + |\underline{v}|_{H^1(T)})$$

due to

$$\|\underline{v}\|_{\hat{T}}^2 \le c_7 h^{-d}\|\underline{v}\|_T^2 \text{ and } \|\underline{v}\|_{H^1(\hat{T})}^2 = \|\underline{v}\|_{\hat{T}}^2 + \|\underline{\nabla}\ \underline{v}\|_{\hat{T}}^2 \le c_7 h^{-d}\|\underline{v}\|_T^2 + c_8 h^{-(d-2)}\|\underline{\nabla}\ \underline{v}\|_T^2$$

Now with the interpolation results for the $H(\text{div})$-conforming interpolation(see e.g. [BF91, III.3.6]) we have

$$\|\underline{v}\|_T = \|\mathbf{\underline{\Pi}}_k^\Sigma \underline{u} - \underline{u}\|_T \le c_{10} h\|\underline{u}\|_{H^1(T)}$$

So finally we get (with $c_{11} = c_5 c_{10}$ and $C = c_{11} + c_4 + c_1$)

$$... \le C \sum_{T\in\mathcal{T}_h} \|\underline{u}\|_{H^1(T)}^2 \le C\|\underline{u}\|_{H^1(\Omega)}^2 x$$

$$\square$$

Proof of Proposition 2.4.3

Proof. We won't proof the Proposition down to all details. We will, at certain points, refer to [KJ98], where the proof for a similar situation has been performed in detail, instead.
We start with representation (2.4.2b) to proof the result. Let's divide volume and facet integrals into two separate expressions for which we will show boundedness:

$$\mathcal{C}_h = (A) + (B) \qquad \text{with}$$

$$(A) = \sum_{T\in\mathcal{T}_h} \int_T \text{div}(\underline{u}\otimes\underline{w})\ \underline{v}\,d\underline{x} = \sum_{T\in\mathcal{T}_h}\int_T (\underline{w}\cdot\underline{\nabla})\underline{u}\ \underline{v}\,d\underline{x} \tag{B.0.1a}$$

$$(B) = \sum_{T\in\mathcal{T}_h}\left\{\int_{\partial T_{in}}|w_n|(\underline{u}^t - \underline{u}_F^t)\ \underline{v}^t\,d\underline{s} + \int_{\partial T_{out}}|w_n|(\underline{u}_F^t - \underline{u}^t)\ \underline{v}_F^t\,d\underline{s}\right\} \tag{B.0.1b}$$

1. There holds

$$(A) \le \|\underline{w}\|_{L^4}\|\underline{\nabla}\ \underline{u}\|_{L^2}\|\underline{v}\|_{L^4} \tag{B.0.2}$$

   by applying Cauchy-Schwarz two times.

2. We now show that $\|\underline{w}\|_{L^4} \le C\|\mathbf{\underline{w}}\|_{1,*}\forall\ w \in [H^1(\mathcal{T}_h)]^d$:

   (a) We pose the problem

   $$-\text{div}(\underline{\nabla}\underline{\phi}) = \underline{w}^3, \underline{\phi} = 0 \text{ on } \partial\Omega \text{ where } (\underline{w}^3)_i = \underline{w}_i^3 \tag{B.0.3}$$

   and as $\underline{w}^3 \in L^4(\Omega)$ we have $\underline{\phi} \in W^{2,p}(\Omega)\cap H_0^1(\Omega)\ \forall\ p \ge 1$ if we assume smooth boundaries.

   (b) The solution operator is continuous onto $W^{2,\frac{4}{3}}(\Omega)$ and so we get

   $$\|\underline{\phi}\|_{2,\frac{4}{3}} \le c_1\|\underline{w}^3\|_{L^{\frac{4}{3}}} = c\|\underline{v}\|_{L^4}^3 \tag{B.0.4}$$

   (c) Now we pose the discrete problem to (B.0.3). Therefore we use a consistent and bounded bilinearform to discretize the vector-valued laplace operator. We use the results of [KJ98, Prop.4.5] where a Standard DG interior penalty bilinearform $\mathcal{X}_h$ is used. Due to consistency there holds

   $$\|\underline{w}\|_{L^4}^4 = (\underline{w}^3, \underline{w}) = \mathcal{X}_h(\underline{\phi}, \underline{w}) \tag{B.0.5}$$

In the proof of [KJ98, Prop.4.5] it was shown that the bilinearform can be bounded in the $\|\cdot\|_{2,\frac{4}{3}}$ norm for $\underline{\phi}$ and in a discrete H^1-seminorm $\|\cdot\|_{1,k}$ similar to one used here for $\underline{w}$:

$$\mathcal{X}_h(\underline{\phi},\underline{w}) \le c_2\|\underline{\phi}\|_{2,\frac{4}{3}}\|\underline{w}\|_{1,k} \tag{B.0.6}$$

With $[\![\underline{w}]\!]^{*2} \le [\![\mathbf{w}^-]\!]^2 + [\![\mathbf{w}^+]\!]^2$ (where $+$ and $-$ indicate the different sides of a facet) the discrete norm $\|\cdot\|_{1,k}$ is smaller than the one used here.

$$\mathcal{X}_h(\underline{\phi},\underline{w}) \le c_2\|\underline{\phi}\|_{2,\frac{4}{3}}\|\mathbf{w}\|_{1,*} \tag{B.0.7}$$

(d) Putting together (B.0.5), (B.0.7) and (B.0.4) we get

$$\|\underline{w}\|_{L^4}^4 \le c_1 c_2 \|\underline{w}\|_{L^4}^3 \|\mathbf{w}\|_{1,*} \Rightarrow \|\underline{w}\|_{L^4} \le c_1 c_2 \|\mathbf{w}\|_{1,*} \tag{B.0.8}$$

3. So we conclude $(A) \le \|\mathbf{w}\|_{1,*}\|\underline{u}\|_{1,*}\|\mathbf{v}\|_{1,*}$

4. We now turn over to the facet integrals.

   (a) We get (with $\|\underline{v}_F^t\| \le \|\underline{v}\| + |[\![\mathbf{v}^t]\!]|$)

$$
\begin{aligned}
(B) \quad \le \quad & \sum_{T\in\mathcal{T}_h} \int_{\partial T} \|\underline{w}\| \; \|[\![\mathbf{u}^t]\!]\| \; (\|\underline{v}\| + |[\![\mathbf{v}^t]\!]|)\, d\underline{s} \\
\le \quad & \sum_{T\in\mathcal{T}_h} h^{\frac{1}{2}} \|\underline{w}\|_{L^4(\partial T)} \left(\|\underline{v}\|_{L^4(\partial T)} + \|[\![\mathbf{v}^t]\!]\|_{L^4(\partial T)} \right) \; \|\underline{u}\|_{1,*}^T \\
\le \quad & \|\mathbf{u}\|_{1,*} \cdot \underbrace{\left(\sum_{T\in\mathcal{T}_h} h\|\underline{w}\|_{L^4(\partial T)}^2 \left(\|\underline{v}\|_{L^4(\partial T)} + \|[\![\mathbf{v}^t]\!]\|_{L^4(\partial T)} \right)^2 \right)^{\frac{1}{2}}}_{(C)}
\end{aligned} \tag{B.0.9}
$$

   (b) And with $a,b \ge 0$ and $(a+b)^2 \le 2(a^2+b^2)$ and $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$ we get

$$(C) \le 2\left(\underbrace{\left(\sum_{T\in\mathcal{T}_h} h\|\underline{w}\|_{L^4(\partial T)}^2\|\underline{v}\|_{L^4(\partial T)}^2 \right)^{\frac{1}{2}}}_{(D)} + \underbrace{\left(\sum_{T\in\mathcal{T}_h} h\|\underline{w}\|_{L^4(\partial T)}^2\|[\![\mathbf{v}^t]\!]\|_{L^4(\partial T)}^2 \right)^{\frac{1}{2}}}_{(E)} \right) \tag{B.0.10}$$

   (c) For (D) the estimates are clear if we use [KJ98, eq. 4.19] in $(*)$:

$$
\begin{aligned}
(D) \quad \le \quad & \left(\sum_{T\in\mathcal{T}_h} h \; \|\underline{w}\|_{L^4(\partial T)}^4 \right)^{\frac{1}{4}} \left(\sum_{T\in\mathcal{T}_h} h \; \|\underline{v}\|_{L^4(\partial T)}^4 \right)^{\frac{1}{4}} \\
\overset{(*)}{\le} \quad & C_{(D)} \left(\|\underline{w}\|_{L^4(\Omega)}^4 + \|\underline{w}\|_{1,k}^4 \right)^{\frac{1}{4}} \left(\|\underline{v}\|_{L^4(\Omega)}^4 + \|\underline{v}\|_{1,k}^4 \right)^{\frac{1}{4}} \\
\le \quad & C_{(D)} \|\mathbf{w}\|_{1,*}\|\mathbf{v}\|_{1,*}
\end{aligned} \tag{B.0.11}
$$

   (d) For (E) the estimate is different and as the hybrid version is less common we have to put some additional effort into this part. Remember that we just consider the two-dimensional case here! We use a scaled version of the Bernstein inequality (see [LS03, p.7]) in $(*)$

which we are allowed to do as $\underline{v}$ is piecewise polynomial. We also use the rough estimate $A_T \geq 0 : \sqrt{\sum_i A_i} \leq \sum_i \sqrt{A_i}$:

$$
\begin{aligned}
(E) \quad &\leq \quad \sum_{T \in \mathcal{T}_h} h^{\frac{1}{2}} \|\underline{w}\|_{L^4(\partial T)} \|[\![\underline{\mathbf{v}}^t]\!]\|_{L^4(\partial T)} \\
&\overset{(*)}{\leq} \quad \sum_{T \in \mathcal{T}_h} c_p h^{\frac{1}{4}} \|\underline{w}\|_{L^4(\partial T)} \|[\![\underline{\mathbf{v}}^t]\!]\|_{L^2(\partial T)} \\
&\leq \quad C \left(\sum_{T \in \mathcal{T}_h} c_p^2 \, h^{\frac{3}{2}} \, \|\underline{w}\|_{L^4(\partial T)}^2 \right)^{\frac{1}{2}} \|\![\underline{\mathbf{v}}]\!\|_{1,*} \\
&\leq \quad C \left(\sum_{T \in \mathcal{T}_h} h \, \|\underline{w}\|_{L^4(\partial T)}^4 \right)^{\frac{1}{4}} \underbrace{\left(\sum_{T \in \mathcal{T}_h} h^2 \, c_p^4 \right)^{\frac{1}{4}}}_{\simeq |\Omega|} \|\![\underline{\mathbf{v}}]\!\|_{1,*} \\
&\leq \quad C_{(E)}(|\Omega|) \, \|\![\underline{\mathbf{w}}]\!\|_{1,*} \|\![\underline{\mathbf{v}}]\!\|_{1,*}
\end{aligned}
\tag{B.0.12}
$$

(e) with $C_{(B)} = 2C_{(D)}C_{(E)}(|\Omega|)$ the estimate holds true for the facet integrals

5. Keep in mind $\|\!\|\cdot\|\!\|_{1,*} \leq \|\!\|\cdot\|\!\|_1$

6. with $\beta_{\mathcal{C}_h}^* = 1 + C_{(B)}$ the proposition holds true.

$\square$

Bibliography

[ABCM02] Douglas N. Arnold, Franco Brezzi, Bernardo Cockburn, L.Donatella Marini
Unified Analysis of Discontinuous Galerkin Methods for Elliptic Problems
SIAM Journal for Numerical Analysis, Vol. 39, No. 5, pp. 1749-1779, 2002

[ARW95] Uri M. Ascher, Steven J. Ruuth, Brian T.R. Wetton
Implicit-explicit methods for time-dependent partial differential equations
SIAM Journal for Numerical Analysis, Vol. 32, No. 3, pp. 797-823, 1995

[ARS97] Uri M. Ascher, Steven J. Ruuth, Raymond J. Spiteri
Implicit-explicit Runge-Kutta methods for time-dependent partial differential equations
Applied Numerical Mathematics, Vol. 25, pp. 151-167, 1997

[Brae97] Dietrich Braess
Finite Elements: theory, fast solvers, and applications in solid mechanics
Cambridge University Press, Cambridge, UK, 1997

[BF91] F. Brezzi, M. Fortin,
Mixed and Hybrid Finite Element Methods
Springer-Verlag, New York, 1991

[BS94] Susanne C. Brenner, L. Ridgway Scott,
The Mathematical Theory of Finite Element Methods
Springer-Verlag, Berlin, 1994

[CGL08] Bernardo Cockburn, Jayadeep Gopalakrishnan, Raytcho Lazarov
Unified Hybridization of Discontinuous Galerkin, Mixed and Continuous Galerkin Methods for Second Order Elliptic Problems
Preprints (http://www.math.ufl.edu/ jayg/research/papers.html, November 2008)

[CKS07] Bernardo Cockburn, Guido Kanschat, Dominik Schötzau
A Note on Discontinuous Galerkin Divergence-free Solutions of the Navier-Stokes Equations
Journal of Scientific Computing Vol. 31, p. 61 ff, May 2007; doi: 10.1007/s10715-006-9107-7

[CKS05] Bernardo Cockburn, Guido Kanschat, Dominik Schötzau,
A locally conservative LDG method for incompressible Navier-Stokes equations
Mathematics of Computations Vol. 74, pp. 1067-1095, 2005

[DB05] Leszek Demkowicz, Annalisa Buffa
H^1,$H(curl)$ and $H(div)$-conforming projection-based interpolation in three dimensions, Quasi-optimal p-interpolation estimates
Comput. Methods Appl. Mech. Eng. 194, pp. 267-296, 2005

[DG10] Leszek Demkowicz, Jayadeep Gopalakrishnan,
A class of Discontinuous Petrov-Galerkin Methods. Part I: The Transport Equation
Computer Methods in Applied Mechanics and Engineering Volume 199, Issues 23-24, pp. 1558-1572,
15 April 2010

[DGS08] Leszek Demkowicz, Jayadeep Gopalakrishnan, Joachim Schöberl
Polynomial Extension Operators. Part I
SIAM J Numerical Analysis, Vol46(6), pp. 3006-3031, 2008

[DH03] Jean Donea and Antonio Huerty
Finite Element Methods for Flow Problems
John Wiley & Sons Ltd, 2003

[ES09] Herbert Egger, Joachim Schöberl
A Mixed-Hybrid-Discontinuous Galerkin Finite Element Method for Convection-Diffusion Problems
IMA Journal of Numerical Analysis 2009; doi: 10.1093/imanum/drn083

[ESW05] Howard Elman, David Silvester, Andy Wathen
Finite Elements and Fast Iterative Solvers with applications in incompressible fluid dynamics
Numerical Mathematics and Scientific Computation, Oxford Science Publications, 2005

[Evan98] Lawrence C. Evans
Partial Differential Equations
Graduate Studies in Mathematics, American Mathematical Society, Providence, Rhode Island, 1998

[EWY02] R. Ewing, J. Wang, and Y. Yang
A Stabilized Discontinuous Finite Element Method for Elliptic Problems
Numerical Linear Algebra with Applications, 2002

[FHV97] J. Frank, W. Hundsdorfer, J.G. Verwer
On the stability of implicit-explicit linear multistep methods
Applied Numerical Mathematics, Vol. 25, pp. 193-205, 1997

[HL02] Peter Hansbo, Mats G. Larson
Discontinuous Galerkin methods for incompressible and nearly incompressible elasticity by Nitsche's method
Comput. Methods Appl. Mech. Eng. 191, pp. 1895-1908, 2002

[KCGH07] Alex Kanevsky, Mark H. Carpenter, David Gottlieb, Jan S. Hesthaven
Application of implicit-explicit high order Runge-Kutta methods to discontinuous-Galerkin schemes
Journal of Computational Physics, Vol. 225, pp. 1753-1781, 2007

[KC03] C.A. Kennedy, Mark H. Carpenter
Additive Runge-Kutta schemes for convection-diffusion-reaction equations
Applied Numerical Mathematics, Vol. 44, pp. 139-181, 2003

[KJ98] Ohannes A. Karakashian, Wadi N. Jureidini
A Nonconforming Finite Element Method for the Stationary Navier-Stokes Equations
SIAM Journal on Numerical Analysis, Vol. 35, pp. 93-120, 1998

[Kova48] L.I.G. Kovasznay
Laminar flow behind a two-dimensional grid
Mathematical Proceedings of the Cambridge Philosophical Society, Vol. 44, pp. 58-62, 1948

[LS03] Andris Lasis, Endre Süli
Poincaré-type inequalities for broken Sobolev Spaces
Technical Report 03/10, Oxford University Computing Laboratory, 2003

[Monk03] Peter Monk
Finite Element Methods for Maxwell's Equations
Numerical Mathematics and Scientific Computation. The Clarendon Press Oxford University Press, New York, 2003

[NPC09] N.C. Nguyen, J.Peraire, B.Cockburn
An implicit high-order hybridizable discontinuous Galerkin method for linear convection-diffusion equations
Journal of Computational Physics, Vol. 228, pp. 3232-3254, 2009

[Schö08] Joachim Schöberl
Divergence-free Hybrid Discontinuous Galerkin Finite Elements for Incompressible Navier Stokes Equations, Augmented Lagrangians, and Robust Preconditioners
Talk, Durham Symposium on CLAPDE, July 2008

[SZ05] Joachim Schöberl, Sabine Zaglmayr
High order Nedelec elements with local complete sequence properties
International Journal for Computation and Mathematics in Electrical and Electronic Engineering, Vol. 24(2), pp. 374-384, 2005

[SZ09] Dominik Schötzau, Liang Zhu
A robust a-posteriori error estimator for discontinuous Galerkin methods for convection-diffusion equations
Applied Numerical Mathematics, Vol. 59, pp. 2236-2255, 2009

[WH03] T. Warburton, J.S. Hesthaven
On the constants in hp-finite element trace inverse inequalities
Comput. Methods Appl. Mech. Eng., Vol. 192, pp. 2765-2773, 2003

[Zagl06] Sabine Zaglmayr
High Order Finite Element Methods for Electromagnetic Field Computation
Dissertation, Institut für Numerische Mathematik, July 2006