Georg-August-Universität Göttingen

Fakultät für Mathematik und Informatik

Institut für Numerische und Angewandte Mathematik

Implicit-Explicit Time Splitting Schemes for Incompressible Navier-Stokes Flows

Masterarbeit

im Studiengang Mathematik (M.Sc.)

eingereicht von: Henry Maximilian von Wahl

eingereicht am: 7. August 2018

Betreuer: Prof. Dr. Gert Lube

Zweitgutachter: Jun.-Prof. Dr. Christoph Lehrenfeld

Acknowledgements

First of all, I would like to thank my supervisor Prof. Dr. Gerd Lube for supervising and supporting me throughout this thesis and teaching me Numerical Mathematics, Partial Differential Equations and Finite Elements over the previous four years in lectures, seminars and other projects. Many thanks also to my second assessor Jun.-Prof. Dr. Christoph Lehrenfeld, with whom I had many insightful discussions and without whom I would not have had the opportunity to work with NGSolve.

I would also like to give a special thanks to Phillip W. Schroeder, whose door was always open for me and who spent many hours helping me in learning to use and work with NGSolve. Many bugs within my code would have remained hidden to me without his patient help.

Furthermore, I would like to thank my father, Leonie Frantzen and Hans-Georg Raumer, with whom I lead many helpful conversations which supported and helped me throughout the time of writing this thesis.

Contents

Introduction

Motivation

Let us consider the unsteady incompressible Navier-Stokes equations [BF13; Joh16; Lay08]

$$\begin{aligned} \partial_t \mathbf{u} - \nu \Delta \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p &= \mathbf{f} \\ \nabla \cdot \mathbf{u} &= 0 \end{aligned} \tag{0.1}$$

in some bounded domain Ω, with Dirichlet boundary conditions $\mathbf{u} = \mathbf{u}_D$ on $\partial\Omega$ and an appropriate initial condition $\mathbf{u}(0) = \mathbf{u}_0$, to model an incompressible Newtonian fluid. Here, $\mathbf{u}$ represents the velocity field, p the pressure, $\mathbf{f}$ an external body force and $\nu = \mathit{const}$ the kinematic viscosity of the fluid under consideration.

We aim to solve (0.1) numerically. From a mathematical aspect, this is an important problem since it is usually "not possible to find an analytical solution to the Navier-Stokes equations" [Joh16]. From a more practical point of view, solving (0.1) numerically is relevant since "wind tunnel tests are very expensive and the conclusions are often uncertain" [Lay08]. One approach to solve the Navier-Stokes equations numerically is the finite element method (FEM), e.g. see [BBF13; DE11; GR86; Joh16]. We will focus focus on *inf-sup stable FEM.* However, classical inf-sup stable FE for instance the class of Taylor-Hood elements present problems such as poor conservation of mass [SL17b] and a lack of pressure robustness [JLM$^+$17].

Pressure robust methods fulfil the fundamental invariance principle that changes in the body force by a gradient field, i.e. $\mathbf{f} \mapsto \mathbf{f} + \nabla\psi$, only affect the pressure and not the velocity, i.e. $(\mathbf{u}, p) \mapsto (\mathbf{u}, p + \psi)$ [JLM$^+$17; SLL$^+$18]. Finite element methods which lack this property lead to error estimates which depend on the pressure estimate multiplied with negative powers of ν [JLM$^+$17; BBF13]. As a result the velocity can be corrupted by poor pressure approximations.

A class of finite elements which are naturally pressure robust are pointwise divergence-free FEM [JLM$^+$17; SLL$^+$18]. These methods give numerical solutions which are pointwise divergence free while using standard, non-divergence free ansatz functions. In the context of $\boldsymbol{\mathcal{H}}^1$-conforming methods, the Scott-Vogelius finite element pair $\mathbb{P}^k/\mathbb{P}^{k-1}_{\text{disc}}$, with some restrictions on the mesh and k [Qua93; Zha04], is an example of such a method. However, in order to have a notion of the divergence, it is not necessary to consider $\boldsymbol{\mathcal{H}}^1$-conforming spaces where we have a notion of the full gradient. In the discontinuous Galerkin (dG) setting, [CKS05] introduced a method where a pointwise divergence free solution is reconstructed in $\boldsymbol{\mathcal{H}}(\text{div}; \Omega)$. Furthermore, a number of $\boldsymbol{\mathcal{H}}(\text{div})$-conforming FEM are also pointwise divergence-free, such as Brezzi-Douglas-Marini and Raviart-Thomas elements with the correct choice of discontinuous pressure space [CKS06; SLL$^+$18]. In order to take advantages of such spatial discretisation, we will use such methods in this thesis for the spatial semi-discretisation of the Navier-Stokes equations.

Spatial discretisation of (0.1) using finite elements leads to differential-algebraic (DAE) system which is non linear in time with a saddle point structure resulting from the velocity-pressure coupling term. To get a discrete solution of the Navier-Stokes equations, we therefore require some time-stepping scheme to solve this DAE. Fully explicit schemes have two major drawbacks. On the one hand, the viscous term is very stiff and therefore stability is only achieved if the chosen time step is very small [HW96; Joh16]. This term should therefore be treated with an implicit scheme. On the other hand, the velocity-pressure coupling term enforces the divergence constraint on the velocity and must therefore also be included in the implicit scheme. However, fully implicit schemes also present some problems. The resulting non-linear system would have to be linearised, e.g. using Newton's method. Here we would have to solve a different, non-symmetric, indefinite system in each step which is computationally expensive [Leh10].

In order to retain the divergence-free nature of our spatial discretisation and avoiding the time step restriction from the viscous term while circumventing the need to solve a non-linear system in each time step, we consider semi-implicit methods. These allow us to treat the Stokes part (viscous and pressure terms) implicitly and the non-linear convective part explicitly.

Semi-implicit schemes which treat the viscous term implicitly and the convective term explicitly have been used for a long time [KM85; KIO91]. Later implicit-explicit (IMEX) multistep schemes were considered in a general ordinary differential equation (ODE) framework by [ARW95] and IMEX Runge-Kutta methods were presented in [ARS97]. Using such IMEX schemes then leads to schemes where the same symmetric system has to be solved in each step with a different right-hand side, which can be implemented efficiently. The price we have to pay for these advantages is a time step restriction, resulting from the explicit treatment of the convective term. However, this restriction is typically less severe than that imposed by an explicitly treated viscous term [Leh10].

IMEX methods have successfully been used in combination with exactly divergence-free FEM, e.g. in [LS16; SJL$^+$18; SLL$^+$18]. In [LS16], a second order IMEX Runge-Kutta method was used while in [SJL$^+$18; SLL$^+$18] an IMEX multistep scheme was used which applies the BDF2 formula to discretise the time-derivative, treats the Stokes part implicitly and splits the convective term to the right-hand side using a second order extrapolation.

Due to our choice of spatial discretisation it is particularly important to us that the IMEX scheme does not destroy the favourable features of the chosen finite element space. However, the methods presented in this thesis are not restricted to our choice of finite element method. They are just as applicable to other discretisation, for example the weakly divergence-free methods such as the Taylor-Hood finite element pair.

The main contribution of this thesis is to derive, collect and analyse IMEX multistep and IMEX Runge-Kutta methods available in the literature which we consider to be suitable to the application of the incompressible Navier-Stokes equations. In particular we asses the schemes with respect to their relative performance in practice.

Outline

This thesis begins in Chapter 1 with the strong formulation of the Navier-Stokes equations in the continuous setting and states relevant functional analytic notation and results. Having introduced the necessary tools we then establish the weak formulation of the Navier-Stokes equations and discuss the solvability thereof.

In Chapter 2 we consider the spatial discretisation of the weak formulation of the Navier-Stokes equations in a $\boldsymbol{\mathcal{H}}(\mathrm{div})$-conforming discontinuous Galerkin setting. To achieve this we introduce some notation from discontinuous Galerkin FEM and derive appropriate weak multilinear forms following [DE11], such as the Symmetric Interior Penalty method for the Laplace problem, in order to discretise the Navier-Stokes equations. The existence and uniqueness of solutions to the semi-discrete problem is then proved and we finish the chapter by introducing $\boldsymbol{\mathcal{H}}(\mathrm{div})$-conforming finite element spaces which lead to pointwise divergence free methods.

With Chapter 3 we begin the temporal discretisation of the DAE resulting from the spatial semi-discretisation discussed before. Following [ARW95], we derive implicit-explicit multistep methods up to order formal 3, i.e. multistep methods which treat the Stokes part implicitly and the non-linear convection part explicitly. The derived schemes are then analysed via a scalar test problem as in [FHV97], and we extend the analysis to the third order method considered here.

As another class of IMEX schemes, we consider IMEX Runge-Kutta schemes in Chapter 4 which were first introduced in [ARS97]. Here the Stokes part is treated with a diagonally implicit Runge-Kutta method while the convection term is treated using a compatible explicit Runge-Kutta method. We give an overview of IMEX Runge-Kutta methods available in the literature which are suitable for the discretisation of the Navier-Stokes equations.

Chapter 5 addresses the practical issues connected with IMEX methods. We discuss the time step restriction resulting from the explicit treatment of the convective term and prove a CFL condition for a scalar transport problem as a test problem for the convective part of the Navier-Stokes equations. We then implement the methods discussed and derived in the earlier chapters with the open source finite element package `NGSolve` using high-order $\boldsymbol{\mathcal{H}}(\mathrm{div})$-conforming finite elements. We consider a variety of problems in both two and three spatial dimensions in order to evaluate how the different IMEX schemes perform in practice with respect to accuracy and computational efficiency. The aim is to identify schemes which are best suited to a given situation.

The results of this thesis are the summarised in Chapter 6, and remaining open problems are discussed. The appendix then covers some implementational details.

1. Preliminaries

In this chapter we will cover the preliminaries for this thesis. We begin by introducing the incompressible Navier-Stokes equations as the model of the movement of an incompressible fluid. We will then cover some basic functional analytic notation and results so that we can bring the Navier-Stokes equations into a weak formulation and finally discuss the solvability thereof.

1.1. The Navier-Stokes Equations

The movement of an incompressible fluid is governed by the time-dependent incompressible Navier-Stokes equations [BF13; Joh16; BIL06] given by

$$\begin{aligned} \partial_t \mathbf{u} - \nu\Delta\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p &= \mathbf{f} && \text{in } (0,T]\times\Omega, \\ \nabla\cdot\mathbf{u} &= 0 && \text{in } (0,T]\times\Omega, \\ \mathbf{u} &= 0 && \text{on } (0,T]\times\partial\Omega, \\ \mathbf{u}(0,\cdot) &= \mathbf{u}_0 && \text{in } \Omega. \end{aligned} \tag{1.1}$$

The domain $\Omega \subset \mathbb{R}^d$ with $d \in \{2,3\}$ is bounded and connected with Lipschitz continuous boundary $\partial\Omega$ which does not depend on time. The function $\mathbf{u} : (0,T]\times\Omega \to \mathbb{R}^d$ represent the velocity field and $p : (0,T]\times\Omega \to \mathbb{R}$ the (zero-mean) kinematic pressure of an incompressible Newtonian fluid moving under the external kinematic body force $\mathbf{f} : (0,T]\times\Omega \to \mathbb{R}^d$ with a suitable initial condition $\mathbf{u}_0$ for the velocity. The constant scalar $\nu = \mu/\rho_0$ is the kinematic viscosity, with μ the dynamic viscosity and ρ_0 the (constant) density of the fluid under consideration.

The first equation in (1.1) represents the balance of momentum, while the second equation enforces the conservation of mass. For ease of presentation, homogeneous Dirichlet boundary conditions have been assumed.

To compare flows on domains of different sizes, we consider the rescaling of the variables in the Navier-Stokes equations into dimensionless variables given by

$$\begin{aligned} \mathbf{x}^* &:= \mathbf{x}/L, \\ \mathbf{u}^* &:= \mathbf{u}/V, \\ t^* &:= Vt/L, \\ p^* &:= [p - p_0]/\rho_0 V^2, \\ \mathbf{f}^* &:= \mathbf{f}L/\rho_0 V^2, \end{aligned}$$

where L is a reference length, V a reference velocity, p_0 a reference pressure and ρ_0 a reference density. Note that with an abuse of notation, here $\mathbf{f}$ and p represent the non-kinematic body forces and pressure respectively. Using the chain rule, the momentum balance equation in the dimensionless variables becomes

$$\partial_{t^*}\mathbf{u}^* - \left(\frac{\mu}{\rho_0 VL}\right)\Delta^*\mathbf{u}^* + (\mathbf{u}^*\cdot\nabla^*)\mathbf{u}^* + \nabla^* p^* = \mathbf{f}^*.$$

The dimensionless parameter $Re := \frac{\rho_0 VL}{\mu} = \frac{VL}{\nu}$ is called the *Reynolds number*. Flows in similar geometries, i.e, if $\Omega^* = \Omega/L$, are then called *dynamically similar* if the Reynolds numbers for the two flows coincide. See [Lay08, Section 5.4] and [Joh16, Section 2.3] for further details.

1.2. Function Spaces

Lebesgue Spaces. Let $\Omega \subset \mathbb{R}^d$ for $d \in \{2,3\}$ be a bounded, connected and Lipschitz continuous domain. For scalar valued functions and $1 \le p \le \infty$ the standard Lebesgue space [DE11; Eva98] is

$$\mathcal{L}^p(\Omega) = \{f : \Omega \to \mathbb{R} \mid f \text{ is measurable and } \|f\|_{\mathcal{L}^p(\Omega)} < \infty\}$$

where the norm is defined by

$$\|v\|_{\mathcal{L}^p(\Omega)} := \begin{cases} \left(\int_\Omega |f|^p \,\mathrm{d}x\right)^{1/p} & \text{for } 1 \le p < \infty, \\ \operatorname{ess\,sup}_{x\in\Omega}|f| & \text{for } p = \infty. \end{cases}$$

We further denote the subspace

$$\mathcal{L}^p_0(\Omega) := \{v \in \mathcal{L}^p(\Omega) \mid \int_\Omega v \,\mathrm{d}x = 0\}$$

of zero-mean $\mathcal{L}^p$ functions. Similarly, we define on the boundary of the domain Ω the space

$$\mathcal{L}^p(\partial\Omega) = \{f : \partial\Omega \to \mathbb{R} \mid f \text{ is measurable and } \|f\|_{\mathcal{L}^p(\partial\Omega)} < \infty\}$$

with the norm

$$\|v\|_{\mathcal{L}^p(\partial\Omega)} := \begin{cases} \left(\int_{\partial\Omega} |f|^p \,\mathrm{d}s\right)^{1/p} & \text{for } 1 \le p < \infty, \\ \operatorname{ess\,sup}_{x\in\partial\Omega}|f| & \text{for } p = \infty. \end{cases}$$

Sobolev Spaces. The standard Sobolev spaces [DE11; Eva98] are then defined for $k \in \mathbb{N}_0$ and $1 \le p \le \infty$ as

$$\mathcal{W}^{k,p}(\Omega) := \{v \in \mathcal{L}^p(\Omega) \mid D^\alpha v \in \mathcal{L}^p(\Omega)\ \forall |\alpha| \le k\}$$

where the derivatives $D^\alpha v$ exist in the weak sense. Equipped with the norm

$$\|v\|_{\mathcal{W}^{k,p}(\Omega)} := \begin{cases} \left(\sum_{|\alpha|\le k} \|v\|^p_{\mathcal{L}^p(\Omega)}\right)^{1/p} & \text{for } 1 \le p < \infty, \\ \sum_{|\alpha|\le k} \|v\|_{\mathcal{L}^\infty(\Omega)} & \text{for } p = \infty, \end{cases}$$

the Sobolev space $\mathcal{W}^{k,p}(\Omega)$ is a Banach space. We will also consider the semi-norm

$$|v|_{\mathcal{W}^{k,p}(\Omega)} := \begin{cases} \left(\sum_{|\alpha|= k} \|v\|^p_{\mathcal{L}^p(\Omega)}\right)^{1/p} & \text{for } 1 \le p < \infty, \\ \sum_{|\alpha|= k} \|v\|_{\mathcal{L}^\infty(\Omega)} & \text{for } p = \infty. \end{cases}$$

The dual space of $\mathcal{W}^{k,p}(\Omega)$ is then denoted by $\mathcal{W}^{-k,p}(\Omega)$. In the case of $p = 2$ we will use the usual notation of $\mathcal{H}^k(\Omega) = \mathcal{W}^{k,2}(\Omega)$. This space equipped with the inner-product

$$(v, w)_{\mathcal{H}^k(\Omega)} := \sum_{|\alpha|\leq k} (D^\alpha v, D^\alpha w)_{\mathcal{L}^2(\Omega)}$$

is a Hilbert space.

Furthermore, we define the so called *Sobolev space with fractional exponent* [EG04] for a general non-integer exponent $0 < s < 1$ and $1 \leq p < \infty$ by

$$\mathcal{W}^{s,p}(\Omega) := \left\{ v \in \mathcal{L}^p(\Omega) \mid \frac{v(x) - v(y)}{\|x - y\|^{s+d/p}} \in \mathcal{L}^p(\Omega \times \Omega) \right\}.$$

For non-integer $s > 1$, we set $\sigma = s - \lfloor s \rfloor$ and define $\mathcal{W}^{s,p}(\Omega)$ by

$$\mathcal{W}^{s,p}(\Omega) := \{ v \in \mathcal{W}^{\lfloor s \rfloor,p}(\Omega) \mid D^\alpha v \in \mathcal{W}^{\sigma,p}(\Omega),\ \forall \alpha,\ |\alpha| = \lfloor s \rfloor \}.$$

As in the integer case for $p = 2$, we use the notation $\mathcal{H}^s(\Omega) = \mathcal{W}^{s,2}(\Omega)$.

Using weak derivatives, we define differential equations in a distributional sense only. In order to make sense of boundary conditions, we need the notion of traces.

Theorem 1.1 (*Trace Theorem for* $\mathcal{H}^s$). *For $r > 1/2$ and $s > 3/2$ there exist surjective trace operators $\gamma_0 : \mathcal{H}^r(\Omega) \to \mathcal{H}^{r-1/2}(\partial\Omega)$ and $\gamma_1 : \mathcal{H}^s(\Omega) \to \mathcal{H}^{s-3/2}(\partial\Omega)$ that are extensions of the boundary values and boundary normal derivatives respectively. Provided that $u \in \mathcal{C}^1(\overline{\Omega})$ it holds that*

$$\gamma_0 u = u|_{\partial\Omega}, \qquad \gamma_1 u = \nabla u \cdot \mathbf{n}|_{\partial\Omega}.$$

Proof. See [Riv08, Section 2.1.3]. □

Now the space $\mathcal{H}^1_0(\Omega)$ denotes the closure of $C^\infty_0(\Omega)$ in $\mathcal{H}^1(\Omega)$. This is the space of $\mathcal{H}^1$-functions whose trace vanishes on the boundary $\partial\Omega$.

Theorem 1.2. *The space $\mathcal{C}^\infty_0(\overline{\Omega})$ is dense in $\mathcal{W}^{k,p}(\Omega)$ with $1 \leq p < \infty$.*

Proof. See [BF13, Theorem III.2.11]. □

Theorem 1.3 (*Generalised Poincaré inequality*). *Let $\Gamma \subset \partial\Omega$ have non-vanishing $(d-1)$-dimensional measure. Then for $1 \leq p < \infty$, there exists $c > 0$ such that for all $u \in \{v \in \mathcal{W}^{1,p}(\Omega) : \gamma_0 v = 0\}$ it holds that*

$$\|u\|_{\mathcal{L}^p(\Omega)} \leq C\|\nabla u\|_{\mathcal{L}^p(\Omega)}.$$

Proof. See [BF13, Proposition III.2.38]. □

Corollary 1.4. *The norm $\|\cdot\|_{\mathcal{H}^1}$ and the semi norm $|\cdot|_{\mathcal{H}^1}$ are equivalent on $\mathcal{H}^1_0(\Omega)$, i.e. $\|\nabla(\cdot)\|_{\mathcal{L}^2(\Omega)}$ is a norm on $\mathcal{H}^1_0(\Omega)$.*

Remark 1.5. We will denote vector-valued spaces by bold-face letters, i.e.

$$\boldsymbol{\mathcal{L}}^2(\Omega) = \left[\mathcal{L}^2(\Omega)\right]^d,\ \boldsymbol{\mathcal{H}}^k(\Omega) = \left[\mathcal{H}^k(\Omega)\right]^d,\ \ldots$$

In the notation of vector- and scalar-valued norms or inner-products we will not distinguish between the two cases. However, the argument of the norm or inner-product will always make it clear which is meant.

Since we will be working with $\boldsymbol{\mathcal{H}}(\mathrm{div})$-conforming FEM we also need the space

$$\boldsymbol{\mathcal{H}}(\mathrm{div};\Omega) \coloneqq \{\mathbf{v} \in \boldsymbol{\mathcal{L}}^2(\Omega) \mid \nabla \cdot \mathbf{v} \in \mathcal{L}^2(\Omega)\}.$$

Together with the inner-product

$$(\mathbf{v}, \mathbf{w})_{\boldsymbol{\mathcal{H}}(\mathrm{div};\Omega)} \coloneqq (\mathbf{v}, \mathbf{w})_{\mathcal{L}^2(\Omega)} + (\nabla \cdot \mathbf{v}, \nabla \cdot \mathbf{w})_{\mathcal{L}^2(\Omega)}$$

this is also a Hilbert space, c.f. [Joh16].

Theorem 1.6. *The space* $\boldsymbol{\mathcal{C}}_0^\infty(\Omega)$ *is dense in* $\boldsymbol{\mathcal{H}}(\mathrm{div};\Omega)$.

Proof. See [BF13, Section IV.3.2]. □

For boundary conditions on the normal component in $\boldsymbol{\mathcal{H}}(\mathrm{div})$ spaces, we again require the notion of traces.

Theorem 1.7 (*Trace Theorem for* $\boldsymbol{\mathcal{H}}^{div}$**).** *There exists a continuous trace operator* $\gamma_n : \boldsymbol{\mathcal{H}}(\mathrm{div};\Omega) \to \mathcal{H}^{-1/2}(\partial\Omega)$ *such that*

$$\gamma_n \mathbf{u} = \mathbf{u} \cdot \mathbf{n}|_{\partial\Omega}$$

holds for all $\mathbf{u} \in \boldsymbol{\mathcal{C}}_0^\infty(\overline{\Omega})$, *where* $\mathbf{n}$ *denotes the outward-pointing unit normal vector on* $\partial\Omega$. *Furthermore, we have the Stokes formula*

$$\int_\Omega \mathbf{u} \cdot \nabla w \,\mathrm{d}\mathbf{x} + \int_\Omega w \nabla \cdot \mathbf{u} \,\mathrm{d}\mathbf{x} = \langle \gamma_n \mathbf{u}, \gamma_0 w \rangle_{\mathcal{H}^{-1/2},\mathcal{H}^{1/2}}$$

for all $\mathbf{u} \in \boldsymbol{\mathcal{H}}(\mathrm{div};\Omega)$ *and* $w \in \mathcal{H}^1(\Omega)$.

Proof. See [BF13, Section IV.3.2] or [EG04, Corollary B.57]. □

Furthermore, we require spaces of divergence-free functions. We define

$$\boldsymbol{\mathcal{C}}_0^{\infty,\mathrm{div}}(\Omega) \coloneqq \{\mathbf{v} \in \boldsymbol{\mathcal{C}}_0^\infty(\Omega) \mid \operatorname{div} \mathbf{v} = 0\}.$$

With $\mathbf{V} = \boldsymbol{\mathcal{H}}_0^1(\Omega)$ we define the space

$$\mathbf{V}^{\mathrm{div}} \coloneqq \{\mathbf{v} \in \mathbf{V} \mid \nabla \cdot \mathbf{v} = 0\}.$$

and similarly, we define

$$\boldsymbol{\mathcal{H}}^{\mathrm{div}} \coloneqq \{\mathbf{v} \in \boldsymbol{\mathcal{H}}(\mathrm{div};\Omega) \mid \nabla \cdot \mathbf{v} = 0, \mathbf{v} \cdot \mathbf{n}|_{\partial\Omega} = 0\}.$$

Bochner Spaces. Since we are considering the *time-dependent* Navier-Stokes equations, we additionally require the concept of *Bochner spaces.*

Definition 1.8. Let $\mathcal{X}$ be a Banach space with norm $\|\cdot\|$. The space

$$\mathcal{L}^p(0,T;\mathcal{X})$$

is the set of all measurable functions $v : [0,T] \to \mathcal{X}$ such that

$$\|v\|_{\mathcal{L}^p(0,T;\mathcal{X})} \coloneqq \begin{cases} \left(\int_0^T \|v(t)\|^p \,\mathrm{d}t\right)^{1/p} < \infty & \text{for } 1 \leq p < \infty, \\ \operatorname{ess\,sup}_{t\in[0,T]} \|v(t)\| < \infty & \text{for } p = \infty. \end{cases}$$

These are the functions which are in $\mathcal{L}^p$ in time and take values in the space $\mathcal{X}$. This space with the above norm is again a Banach space, c.f. [BF13, Proposition II.5.2].

1.3. The Ladyzhenskaya-Babuska-Brezzi Condition

Following [GR86; Riv08], let us consider following abstract setting: Let X and Q be Hilbert spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Q$ respectively. Denote by X' and Q' the dual spaces of X and Q respectively. Then consider a continuous bilinear form $b(\cdot,\cdot) : X \times Q \to \mathbb{R}$ and define the operator

$$B : X \to Q'$$

and its dual operator

$$B' : Q \to X'$$

by

$$Bv(q) = B'q(v) = b(v,q).$$

We then define the kernel of B:

$$V = \ker(B) = \{v \in X \mid b(v,q) = 0 \quad \forall q \in Q\}$$

as well as its orthogonal set

$$V^{\perp} = \{w \in X \mid (v,w)_X = 0 \quad \forall v \in V\}$$

and its polar set

$$V^{\circ} = \{\phi \in X' \mid \phi(v) = 0 \quad \forall v \in V\}.$$

We can then formulate the following result:

Lemma 1.9 (*The inf-sup Condition*). *The following properties are equivalent*

(i) The Ladyzhenskaya-Babuška-Brezzi condition is satisfied, i.e. there exists $\beta > 0$ such that

$$\inf_{q\in Q\setminus\{0\}} \sup_{v\in X\setminus\{0\}} \frac{b(v,q)}{\|q\|_Q \|v\|_X} > \beta. \tag{LBB}$$

(ii) The operator B is an isomorphism from $V^{\perp}$ onto Q' and

$$\|Bv\|_{Q'} \geq \beta \|v\|_X \,.$$

(iii) The dual operator B' is an isomorphism from Q onto V° and

$$\|B'q\|_{X'} \geq \beta \|q\|_Q \,.$$

Proof. See [GR86, Lemma 4.1]. □

Lemma 1.10. *Let Ω be a bounded and Lipschitz continuous domain in $\mathbb{R}^d$. Then the pair $X = \boldsymbol{\mathcal{H}}_0^1(\Omega)$ and $Q = \mathcal{L}_0^2(\Omega)$ satisfy the condition* (LBB) *with $b(\mathbf{v},q) = \int_\Omega p\nabla\cdot\mathbf{v}\,\mathrm{d}\mathbf{x}$.*

Proof. See [EG04, Corollary B.71] or [Riv08, Lemma 6.4]. □

Remark 1.11. Pairs of spaces that fulfil the Ladyzhenskaya-Babuška-Brezzi condition will be referred to as *inf-sup stable.*

1.4. The Weak Formulation of the Navier-Stokes Equations

Let $\mathbf{v} \in \boldsymbol{\mathcal{C}}_0^\infty(\Omega)$ be an arbitrary test function. Multiplying the momentum balance equation of the Navier-Stokes equations (1.1) with $\mathbf{v}$, integrating and using integration by parts for the viscous and pressure terms gives

$$\partial_t \int_\Omega \mathbf{u} \cdot \mathbf{v} \,\mathrm{d}\mathbf{x} + \nu \int_\Omega \nabla \mathbf{u} : \nabla \mathbf{v} \,\mathrm{d}\mathbf{x} + \int_\Omega ((\mathbf{u} \cdot \nabla)\mathbf{u}) \cdot \mathbf{v} \,\mathrm{d}\mathbf{x} - \int_\Omega p \operatorname{div} \mathbf{v} \,\mathrm{d}\mathbf{x} = \int_\Omega \mathbf{f} \cdot \mathbf{v} \,\mathrm{d}\mathbf{x} \tag{1.2}$$

where the boundary integrals vanish since $\mathbf{v}|_{\partial\Omega} = 0$. Due to the density of $\boldsymbol{\mathcal{C}}_0^\infty(\Omega)$ in $\boldsymbol{\mathcal{H}}_0^1(\Omega)$, we may consider test functions $\mathbf{v} \in \mathbf{V} = \boldsymbol{\mathcal{H}}_0^1(\Omega)$. To remove the pressure and include the conservation of mass, we simplify the problem by restricting the problem to $\mathbf{v} \in \mathbf{V}^{\mathrm{div}}$. We then have the following problem:

Problem P1 (*Weak Formulation - Time-independent test functions*). For $\mathbf{u}_0 \in \boldsymbol{\mathcal{H}}^{\mathrm{div}}$ and $\mathbf{f} \in \mathcal{L}^2(0,T;\mathbf{V}')$ find $\mathbf{u} \in \mathcal{L}^2(0,T;\mathbf{V}^{\mathrm{div}})$ such that

$$\begin{aligned} \partial_t \int_\Omega \mathbf{u}(t) \cdot \mathbf{v} \,\mathrm{d}\mathbf{x} + \nu \int_\Omega \nabla \mathbf{u}(t) : \nabla \mathbf{v} \,\mathrm{d}\mathbf{x} + \int_\Omega ((\mathbf{u}(t) \cdot \nabla)\mathbf{u}(t)) \cdot \mathbf{v} \,\mathrm{d}\mathbf{x} &= \langle f(t), \mathbf{v} \rangle_{V',V} \\ \mathbf{u}(0) &= \mathbf{u}_0 \end{aligned} \tag{1.3}$$

for all $\mathbf{v} \in \mathbf{V}^{\mathrm{div}}$ in $(\boldsymbol{\mathcal{C}}_0^\infty((0,T)))'$.

Remark 1.12. A solution of this problem is weakly continuous from $[0,T]$ into $\boldsymbol{\mathcal{H}}^{\mathrm{div}}$. Therefore the initial condition makes sense, even though $\mathbf{u}$ is only in $\mathcal{L}^2([0,T])$ [Tem77].

This definition of the weak formulation is insufficient as it is necessary to test with the solution to get the energy equation. To solve this we use time-dependent test functions. Let therefore $\mathbf{v} \in \mathcal{C}_0^\infty(0,T;\boldsymbol{\mathcal{C}}_0^{\infty,\mathrm{div}}(\Omega))$. We multiply the momentum balance equation with $\mathbf{v}$, integrate with respect to Ω and use integration by parts for the diffusion term. Integrating with respect to time and using integration by parts to shift the time-derivative onto the test function then gives

$$\begin{aligned} -\int_0^T \int_\Omega \mathbf{u} \cdot \partial_t \mathbf{v} \,\mathrm{d}\mathbf{x} \,\mathrm{d}t + \int_0^T \left[\nu \int_\Omega \nabla \mathbf{u}(t) : \nabla \mathbf{v} \,\mathrm{d}\mathbf{x} + \int_\Omega ((\mathbf{u}(t) \cdot \nabla)\mathbf{u}(t)) \cdot \mathbf{v} \,\mathrm{d}\mathbf{x} \right] \mathrm{d}t& \\ = \int_0^T \langle f(t), \mathbf{v} \rangle_{V',V} \,\mathrm{d}t& \end{aligned} \tag{1.4}$$

where space integrals at $t \in \{0,T\}$ vanish due to $\mathbf{v}(0) = \mathbf{v}(T) = 0$ in Ω.

Problem P2 (*Weak Formulation - Time-dependent test functions*). For $\mathbf{u}_0 \in \boldsymbol{\mathcal{H}}^{\mathrm{div}}$ and $\mathbf{f} \in \mathcal{L}^2(0,T;\mathbf{V}')$ find $\mathbf{u} \in \mathcal{L}^2(0,T;\mathbf{V}^{\mathrm{div}})$ such that $\partial_t \mathbf{u} \in \mathcal{L}^1(0,T;(\mathbf{V}^{\mathrm{div}})')$ which satisfies (1.4) for all $\mathbf{v} \in \mathcal{C}_0^\infty(0,T;\mathbf{V}^{\mathrm{div}})$ and $\mathbf{u}(0) = \mathbf{u}_0$.

Remark 1.13. Here the initial condition for the solution immediately makes sense as functions $\mathbf{u} \in \mathcal{L}^2(0,T;\mathbf{V}^{\mathrm{div}})$ for which additionally $\partial_t \mathbf{u} \in \mathcal{L}^1(0,T;(\mathbf{V}^{\mathrm{div}})')$ holds, are continuous with values in $(\mathbf{V}^{\mathrm{div}})'$ in the strong topology [BF13].

Theorem 1.14 (*Equivalence of Weak Formulations*). *Let* $\mathbf{f} \in \mathcal{L}^1(0,T;(\mathbf{V}^{div})')$ *and let* $\mathbf{u} \in \mathcal{L}^2(0,T;\mathbf{V}^{div})$*. Then* $\mathbf{u}$ *has a weak derivative* $\partial_t \mathbf{u} \in \mathcal{L}^1(0,T;(\mathbf{V}^{div})')$ *and satisfies* (1.4) *iff* $\mathbf{u}$ *satisfies* (1.3)*.*

Proof. See [BF13, Chapter V]. □

Theorem 1.15 (*Solvability*). *Let Ω be a bounded and connected domain in $\mathbb{R}^d$ with Lipschitz continuous boundary $\partial\Omega$ and let $\nu > 0$ be given. For $\mathbf{u}_0 \in \boldsymbol{\mathcal{H}}^{div}$ and $\mathbf{f} \in \mathcal{L}^2(0,T;\boldsymbol{\mathcal{H}}^{-1})$ there exists a weak solution of the Navier-Stokes equations*

$$(\mathbf{u}, p) \in \Big(\mathcal{L}^2(0,T;\mathbf{V}^{div}) \cap \mathcal{L}^\infty(0,T;\boldsymbol{\mathcal{H}}^{div})\Big) \times \mathcal{W}^{-1,\infty}(0,T;\mathcal{L}^2_0(\Omega))$$

with

$$\partial_t \mathbf{u} \in \mathcal{L}^{4/d}(0,T;\mathbf{V}^{div})$$

satisfying

$$\frac{1}{2}\|\mathbf{u}(t)\|^2_{\mathcal{L}^2(\Omega)} + \nu\int_0^t \|\nabla\mathbf{u}(\tau)\|^2_{\mathcal{L}^2(\Omega)}\,\mathrm{d}\tau \le \frac{1}{2}\|\mathbf{u}_0\|^2_{\mathcal{L}^2(\Omega)} + \int_0^t \langle\mathbf{f}(\tau),\mathbf{u}(\tau)\rangle_{V',V}\,\mathrm{d}\tau \tag{1.5}$$

for all $t \in [0,T]$. In the case $d = 2$, the solution is unique and we have equality in (1.5).

Proof. See [Gal00] or [BF13, Chapter V]. □

Remark 1.16. The uniqueness of such a solution in three dimensions is still an open problem. It is possible to prove uniqueness of a solution $\mathbf{u} \in \mathcal{L}^s(0,T;\boldsymbol{\mathcal{L}}^q(\Omega))$ with $s > 2, q > 3$ and ${}^2/_s + {}^3/_q = 1$, however, the existence of such a solution is unknown [Joh16]. It is also an open question whether "this gap reflects an essential property of real fluids or an inadequacy in the model or analysis" [Lay08]. Connected to this gap is the *Millennium Problem* of the Clay Mathematics Institute as described in [Fef00].

2. Spatial Discretisation

We will now consider the spatial discretisation of the weak formulation of the incompressible Navier-Stokes equations in an $\boldsymbol{\mathcal{H}}(\mathrm{div})$-conforming finite element context. To this end we begin by introducing some discontinuous Galerkin FEM notation. With this we then derive discrete forms for the corresponding continuous weak multilinear forms for the viscous, convective and pressure-velocity coupling terms respectively. With the discrete problem formulated, we will discuss the solvability thereof and under certain assumptions state a discretisation error estimate and convergence result from the literature. Finally, we give some examples of $\boldsymbol{\mathcal{H}}(\mathrm{div})$-conforming finite element spaces.

2.1. Discrete Setting

2.1.1. Notation

Let $\mathcal{T}_h$ be a shape-regular mesh of the domain Ω with mesh size

$$h \coloneqq \max_{T \in \mathcal{T}_h} h_T$$

where h_T denotes the diameter of the element $T \in \mathcal{T}_h$. For simplicity, we assume that $\partial\Omega$ is polygonal in order for the mesh to describe the domain exactly. Since we will be working with dG-FEM we require the concept of mesh facets, averages and jumps. For further details we refer to [DE11; Riv08].

Definition 2.1. Let $\mathcal{T}_h$ be a mesh of $\Omega \subset \mathbb{R}^d$. A (closed) subset F of $\overline{\Omega}$ is called a *mesh facet* if F has positive $(d-1)$-dimensional Hausdorff measure and if one of the following conditions holds:

1. There exist mesh elements $T_1 \neq T_2$ such $F = \partial T_1 \cap \partial T_2$. In this case F is called an *interface*.
2. There exists a mesh element T such that $F = \partial T \cap \partial\Omega$. In this case F is called a *boundary facet*.

For a given mesh $\mathcal{T}_h$, the set of interfaces is the denoted by $\mathcal{F}_h^i$ and the set of boundary facets is denoted by $\mathcal{F}_h^b$. We then denote the complete set of facets by

$$\mathcal{F}_h \coloneqq \mathcal{F}_h^i \cup \mathcal{F}_h^b.$$

Furthermore, the set

$$\mathcal{F}_T \coloneqq \{F \in \mathcal{F}_h \mid F \subset \partial T\}$$

contains the facets which are part of the boundary of a single element $T \in \mathcal{T}_h$. The maximum number of facets composing the boundary the mesh elements is denoted by

$$N_\partial \coloneqq \max_{T \in \mathcal{T}_h} \operatorname{card}(\mathcal{F}_T). \tag{2.1}$$

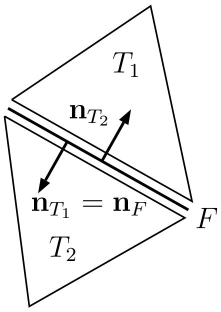


Figure 2.1.: Unit normal vectors on an interior facet in the dG setting.

Definition 2.2. Let v be a scalar-valued function which is sufficiently smooth to admit a (possibly two valued) trace on all interfaces. Then for all $F \in \mathcal{F}_h^i$ we define the *jump*

$$[\![v]\!]_F := v|_{T_1} - v|_{T_2}$$

and the *average*

$$\{\!\{v\}\!\}_F := \frac{1}{2}\left(v|_{T_1} + v|_{T_2}\right).$$

If v is vector valued, we define the jump and average to act component-wise. Furthermore, we will drop the index F whenever no confusion can arise.

Remark 2.3. We extend the notion of jumps and averages to boundary facets. For all $F \in \mathcal{F}_h^b$ we set

$$[\![v]\!]_F = \{\!\{v\}\!\}_F = v|_T.$$

Lemma 2.4. *For two sufficiently smooth functions u and v which admit a possibly two valued trace on all interfaces, we have*

$$[\![uv]\!] = [\![u]\!]\{\!\{v\}\!\} + \{\!\{u\}\!\}[\![v]\!].$$

Proof. On every interface $F = \partial T_1 \cap \partial T_2$ we have

$$\begin{aligned}[\![uv]\!] &= u|_{T_1} v|_{T_1} - u|_{T_2} v|_{T_2} \\ &= \frac{1}{2}\left(u|_{T_1} + u|_{T_2}\right)\left(v|_{T_1} - v|_{T_2}\right) + \left(u|_{T_1} - u|_{T_2}\right)\frac{1}{2}\left(v|_{T_1} + v|_{T_2}\right) \\ &= \{\!\{u\}\!\}[\![v]\!] + [\![u]\!]\{\!\{v\}\!\}.\end{aligned}$$

□

Definition 2.5. For all $F \in \mathcal{F}_h$ we define the *unit normal* $\mathbf{n}_F$ to F as

1. $\mathbf{n}_{T_1}$, the unit normal vector pointing from T_1 to T_2 if $\partial T_1 \cap \partial T_2 = F \in \mathcal{F}_h^i$.
2. $\mathbf{n}$, the unit outward pointing normal to Ω if $F \in \mathcal{F}_h^b$.

The orientation of the unit normal on an interface is arbitrary depending on the choice of T_1 and T_2, however, it is kept fixed for the remainder of this thesis. See Figure 2.1.

2.1.2. Broken Sobolev Spaces

For a given mesh $\mathcal{T}_h$ and $k \in \mathbb{N}_0$ we define the *broken Sobolev space* as

$$\mathcal{H}^k(\mathcal{T}_h) \coloneqq \{v \in \mathcal{L}^2(\Omega) \mid v|_T \in \mathcal{H}^k(T) \text{ for all } T \in \mathcal{T}_h\}.$$

Definition 2.6. The *broken gradient* $\nabla_h : \mathcal{H}^1(\mathcal{T}_h) \to \boldsymbol{\mathcal{L}}^2(\Omega)$ is defined such that

$$(\nabla_h v)\big|_T \coloneqq \nabla(v|_T)$$

for all $T \in \mathcal{T}_h$. For simplicity, we therefore drop the index h in case we consider the broken gradient restricted to a single fixed element.

Lemma 2.7. *For all $v \in \mathcal{H}^1(\Omega)$ we have $\nabla_h v = \nabla v$ in $\boldsymbol{\mathcal{L}}^2(\Omega)$.*

Proof. C.f. [DE11, Lemma 1.22]. □

We also define the broken $\boldsymbol{\mathcal{H}}(\mathrm{div})$ space

$$\boldsymbol{\mathcal{H}}(\mathrm{div}; \mathcal{T}_h) \coloneqq \{\mathbf{v} \in \boldsymbol{\mathcal{L}}^2(\Omega) \mid v|_T \in \boldsymbol{\mathcal{H}}(\mathrm{div}; T) \text{ for all } T \in \mathcal{T}_h\}.$$

Definition 2.8. The *broken divergence operator* $\nabla_h \cdot : \boldsymbol{\mathcal{H}}(\mathrm{div}; \mathcal{T}_h) \to \mathcal{L}^2(\Omega)$ is defined such that

$$(\nabla_h \cdot \mathbf{v})\big|_T \coloneqq \nabla \cdot (\mathbf{v}|_T).$$

Lemma 2.9. *For all $\mathbf{v} \in \boldsymbol{\mathcal{H}}(\mathrm{div}; \Omega)$ we have $\nabla_h \cdot \mathbf{v} = \nabla \cdot \mathbf{v}$ in $\mathcal{L}^2(\Omega)$.*

Proof. C.f. [DE11, Section 1.2.6]. □

Lemma 2.10 (*Characterisation of $\boldsymbol{\mathcal{H}}(\mathrm{div}; \Omega)$, c.f. [DE11; Joh16]*). *Let $\mathcal{T}_h$ be an admissible triangulation of the domain Ω. For a function $\mathbf{w} \in \boldsymbol{\mathcal{H}}(\mathrm{div}; \mathcal{T}_h) \cap \mathcal{W}^{1,1}(\mathcal{T}_h)$ we have that $\mathbf{w} \in \boldsymbol{\mathcal{H}}(\mathrm{div}; \Omega)$ if and only if*

$$[\![\mathbf{w}]\!] \cdot \mathbf{n}_F = 0 \qquad \textit{for all } F \in \mathcal{F}_h^i. \tag{2.2}$$

Proof. Let $\mathbf{w} \in \boldsymbol{\mathcal{H}}(\mathrm{div}; \mathcal{T}_h) \cap \mathcal{W}^{1,1}(\mathcal{T}_h)$ and $\varphi \in C_0^\infty(\Omega)$. Integration by parts then yields

$$\begin{aligned}
\int_\Omega \mathbf{w} \cdot \nabla\varphi \,\mathrm{d}\mathbf{x} &= \sum_{T \in \mathcal{T}_h} \int_T \mathbf{w} \cdot \nabla\varphi \,\mathrm{d}\mathbf{x} \\
&= -\sum_{T \in \mathcal{T}_h} \left(\int_T (\nabla \cdot \mathbf{w})\varphi \,\mathrm{d}\mathbf{x} - \int_{\partial T} (\mathbf{w} \cdot \mathbf{n}_T)\varphi \,\mathrm{d}\mathbf{s} \right) \\
&= -\sum_{T \in \mathcal{T}_h} \int_T (\nabla \cdot \mathbf{w})\varphi \,\mathrm{d}\mathbf{x} + \sum_{T \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_T} \int_F (\mathbf{w} \cdot \mathbf{n}_F)\varphi \,\mathrm{d}\mathbf{s} \\
&= -\int_\Omega (\nabla_h \cdot \mathbf{w})\varphi \,\mathrm{d}\mathbf{x} + \sum_{F \in \mathcal{F}_h^i} \int_F ([\![\mathbf{w}]\!] \cdot \mathbf{n}_F)\varphi \,\mathrm{d}\mathbf{s} + \sum_{F \in \mathcal{F}_h^b} \int_F (\mathbf{w} \cdot \mathbf{n}_F)\varphi \,\mathrm{d}\mathbf{s} \qquad (2.3)
\end{aligned}$$

where the last term in (2.3) vanishes since $\varphi|_{\partial\Omega} = 0$. Therefore, if we have $[\![\mathbf{w}]\!] \cdot \mathbf{n}_F = 0$ for all $F \in \mathcal{F}_h^i$ it holds

$$\int_\Omega \mathbf{w} \cdot \nabla\varphi \,\mathrm{d}\mathbf{x} = -\int_\Omega (\nabla_h \cdot \mathbf{w})\varphi \,\mathrm{d}\mathbf{x} \qquad \text{for all } \varphi \in C_0^\infty(\Omega). \tag{2.4}$$

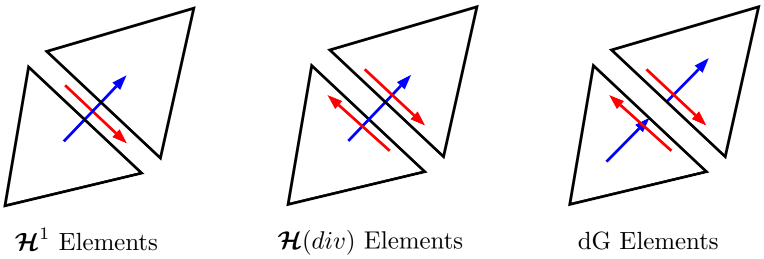


Figure 2.2.: (Dis-)continuity of the tangential component (red) and the normal component (blue) for different FEM.

This means that $\nabla \cdot \mathbf{w} = \nabla_h \cdot \mathbf{w} \in \mathcal{L}^2(\Omega)$. Conversely, if (2.4) holds then by the above computation it follows that

$$\sum_{F \in \mathcal{F}_h^i} \int_F ([\![\mathbf{w}]\!] \cdot \mathbf{n}_T)\varphi \, \mathrm{d}\mathbf{s} = 0.$$

We then get (2.2) by choosing the support of φ to intersect exactly one interface. This is possible since φ is arbitrary in $C_0^\infty(\Omega)$. □

Figure 2.2 illustrates this result by comparing the continuity of the tangential and normal components across element facets for different finite elements.

2.2. Discrete Problem

Let us assume that we have sufficiently regular data and a sufficiently regular exact solution of the Navier-Stokes equation so that we can consider the weak formulation (1.3). To discretise this, we split the problem into three parts and consider the diffusion part

$$a(\mathbf{u}, \mathbf{v}) = \int_\Omega \nabla \mathbf{u} : \nabla \mathbf{v} \, \mathrm{d}\mathbf{x}$$

the velocity pressure coupling

$$b(\mathbf{v}, q) = -\int_\Omega q \nabla \cdot \mathbf{v} \, \mathrm{d}\mathbf{x}$$

and the convection term

$$c(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \int_\Omega ((\mathbf{u} \cdot \nabla)\mathbf{v}) \cdot \mathbf{w} \, \mathrm{d}\mathbf{x}$$

separately.

Discrete Spaces. To keep this section as general as possible, we follow [SL17a; SLL$^+$18] and consider a general FEM velocity-pressure space pair $(\mathbf{V}_h, Q_h)$ on which we place the following assumptions:

Assumption A1.

$$\mathbf{V}_h = \{\mathbf{v} \in \mathcal{H}(\mathrm{div}; \Omega) \ : \ \mathbf{v}|_T \in \mathbf{V}^k(T), \ \forall T \in \mathcal{T}_h; \ \mathbf{v} \cdot \mathbf{n}|_{\partial\Omega} = 0\} \subset \mathcal{H}(\mathrm{div}; \Omega)$$
$$Q_h = \{q \in \mathcal{L}_0^2(\Omega) \ : \ q|_T \in \mathbb{P}^l(T), \ \forall T \in \mathcal{T}_h\} \subset \mathcal{L}_0^2(\Omega)$$

with $l \in \{k-1, k\}$ and a local space $\mathbf{V}^k(T)$ of polynomial order k. The choice of local space will be specified in Section 2.3. Note that this also includes $\mathcal{H}^1$-conforming FEM but excludes fully discontinuous dG-FEM due to Lemma 2.10. However, we will only focus on $\boldsymbol{\mathcal{H}}(\mathrm{div})$-conforming methods.

Assumption A2. The spaces $\mathbf{V}_h$ and Q_h are *divergence conforming*, i.e.

$$\nabla \cdot \mathbf{V}_h \subseteq Q_h. \tag{2.5}$$

This condition ensures the pointwise divergence free nature of the velocity solution as we will show in Lemma 2.18.

Assumption A3. The FEM spaces $\mathbf{V}_h$ and Q_h fulfil the *discrete Babuška-Brezzi condition*, i.e. there exists $\beta_0 > 0$, independent of h and k, such that

$$\inf_{q_h \in Q_h \setminus \{0\}} \sup_{\mathbf{v} \in \mathbf{V}_h \setminus \{0\}} \frac{b(\mathbf{v}_h, q_h)}{\|q_h\|_{Q_h} \|\mathbf{v}_h\|_{V_h}} \geq \beta_h > \beta_0 > 0. \tag{LBB$_h$}$$

Analogue to the smooth case we then say that $\mathbf{V}_h$ and Q_h form an *inf-sup* stable FE pair. We will see in Section 2.2.2 that in the $\boldsymbol{\mathcal{H}}(\mathrm{div})$-conforming FE context the continuous velocity pressure coupling term $b(\cdot,\cdot)$ and the discrete counterpart $b_h(\cdot,\cdot)$ coincide.

In Section 2.3 we will go into further detail of FEM pairs which fulfil these assumptions and which will be used in numerical experiments conducted later.

2.2.1. The Symmetric Interior Penalty Method for Diffusion

As a test problem for diffusion, we begin with the weak formulation of the Poisson problem with homogeneous Dirichlet boundary conditions

$$\text{Find } \mathbf{u} \in \mathbf{V} \text{ such that } a(\mathbf{u}, \mathbf{v}) = f(\mathbf{v}) \text{ for all } \mathbf{v} \in \mathbf{V} \tag{2.6}$$

with $\mathbf{V} = \boldsymbol{\mathcal{H}}_0^1(\Omega)$, the bilinear form $a(\mathbf{u}, \mathbf{v}) = \int_\Omega \nabla \mathbf{u} : \nabla \mathbf{v} \, d\mathbf{x}$ and the linear form $f(\mathbf{v}) = \int_\Omega \mathbf{f} \cdot \mathbf{v} \, d\mathbf{x}$. Following [DE11], we want to compute a discrete approximate of the solution to (2.6) in the broken space $\mathbf{V}_h \subset \mathbf{V}$ from the discrete problem

$$\text{Find } \mathbf{u}_h \in \mathbf{V}_h \text{ such that } a_h(\mathbf{u}_h, \mathbf{v}) = f_h(\mathbf{v}) \text{ for all } \mathbf{v} \in \mathbf{V}_h \tag{2.7}$$

where $a_h : \mathbf{V}_h \times \mathbf{V}_h \to \mathbb{R}$ and $f_h : \mathbf{V}_h \to \mathbb{R}$ are some appropriate discrete bilinear and linear forms approximating $a(\cdot,\cdot)$ and $f(\cdot)$.

Consistency. We assume, that there exists a subspace $\mathbf{V}^* \subset \mathbf{V}$ so that the exact solution $\mathbf{u}$ is in $\mathbf{V}^*$ and that we can extend the discrete bilinear form to $\mathbf{V}^* \times \mathbf{V}_h$ such that we are able to plug in the exact solution into the first argument. Note that an extension to $\mathbf{V} \times \mathbf{V}_h$ is in general not possible [DE11]. In order to ensure that all terms during the derivation remain well posed we define

$$\mathbf{V}^* = \mathbf{V}_h \oplus (\mathbf{V} \cap \boldsymbol{\mathcal{H}}^2(\mathcal{T}_h))$$

on which we derive the SIP bilinear form. Note, however, that we will see that the final bilinear form will be well-posed for

$$\mathbf{V}^* = \mathbf{V}_h \oplus (\mathbf{V} \cap \boldsymbol{\mathcal{H}}^s(\mathcal{T}_h))$$

with some $s > 3/2$ due to Theorem 1.1 and the appearance of normal derivatives on facets.

Definition 2.11. We call the discrete problem (2.7) *consistent* if for the exact solution $\mathbf{u} \in \mathbf{V}^*$ of (2.6) we have

$$a_h(\mathbf{u}, \mathbf{v}_h) = f_h(\mathbf{v}_h) \qquad \text{for all } \mathbf{v}_h \in \mathbf{V}_h. \tag{2.8}$$

Remark 2.12. We observe that the definition of consistency is equivalent to *Galerkin orthogonality*, i.e, we have that (2.8) holds if and only if

$$a_h(\mathbf{u} - \mathbf{u}_h, \mathbf{v}_h) = 0 \qquad \text{for all } \mathbf{v}_h \in \mathbf{V}_h.$$

Derivation. Let $\mathcal{T}_h$ be an admissible triangulation of the domain Ω. To derive the symmetric interior penalty formulation for the bilinear form $a(\mathbf{u}, \mathbf{v}) = \int_\Omega \nabla \mathbf{u} : \nabla \mathbf{v} \, \mathrm{d}\mathbf{x}$ we begin by localising to each element and applying integration by parts

$$\begin{aligned} a_h^{(0)}(\mathbf{v}, \mathbf{w}) &= \int_\Omega \nabla_h \mathbf{v} : \nabla_h \mathbf{w} \, \mathrm{d}\mathbf{x} \\ &= \sum_{T \in \mathcal{T}_h} \int_T \nabla \mathbf{v} : \nabla \mathbf{w} \, \mathrm{d}\mathbf{x} \\ &= - \sum_{T \in \mathcal{T}_h} \int_T (\Delta \mathbf{v}) \mathbf{w} \, \mathrm{d}\mathbf{x} + \sum_{T \in \mathcal{T}_h} \int_{\partial T} (\nabla \mathbf{v} \cdot \mathbf{n}_T) \cdot \mathbf{w} \, \mathrm{d}\mathbf{s}. \end{aligned}$$

Observing that for all interior facets $F = \partial T_1 \cap \partial T_2$ we have $\mathbf{n}_F = \mathbf{n}_{T_1} = -\mathbf{n}_{T_2}$, we may rewrite the second term as a sum over facets

$$\begin{aligned} \sum_{T \in \mathcal{T}_h} \int_{\partial T} (\nabla \mathbf{v} \cdot \mathbf{n}_T) \cdot \mathbf{w} \, \mathrm{d}\mathbf{s} = & \sum_{F \in \mathcal{F}_h^i} \int_F (\nabla \mathbf{v}|_{T_1} \cdot \mathbf{n}_{T_1}) \cdot \mathbf{w}|_{T_1} + (\nabla \mathbf{v}|_{T_2} \cdot \mathbf{n}_{T_2}) \cdot \mathbf{w}|_{T_2} \, \mathrm{d}\mathbf{s} \\ & + \sum_{F \in \mathcal{F}_h^b} \int_F (\nabla \mathbf{v} \cdot \mathbf{n}_F) \cdot \mathbf{w} \, \mathrm{d}\mathbf{s} \\ = & \sum_{F \in \mathcal{F}_h^i} \int_F [\![(\nabla \mathbf{v}) \mathbf{w}]\!] \cdot \mathbf{n}_F \, \mathrm{d}\mathbf{s} + \sum_{F \in \mathcal{F}_h^b} \int_F (\nabla \mathbf{v} \cdot \mathbf{n}_F) \cdot \mathbf{w} \, \mathrm{d}\mathbf{s}. \end{aligned}$$

Now using Lemma 2.4 we have

$$[\![(\nabla_h \mathbf{v}) \mathbf{w}]\!] = \{\!\{ \nabla_h \mathbf{v} \}\!\} [\![\mathbf{w}]\!] + [\![\nabla_h \mathbf{v}]\!] \{\!\{ \mathbf{w} \}\!\}.$$

Taking into account the definition of jumps and averages on boundary facets we obtain

$$a_h^{(0)}(\mathbf{v}, \mathbf{w}) = - \sum_{T \in \mathcal{T}_h} \int_T (\Delta \mathbf{v}) \mathbf{w} \, \mathrm{d}\mathbf{x} + \sum_{F \in \mathcal{F}_h} \{\!\{ \nabla_h \mathbf{v} \}\!\} \cdot \mathbf{n}_F \cdot [\![\mathbf{w}]\!] + \sum_{F \in \mathcal{F}_h^i} [\![\nabla_h \mathbf{v}]\!] \cdot \mathbf{n}_F \cdot \{\!\{ \mathbf{w} \}\!\}.$$

Inserting the exact solution $\mathbf{v} = \mathbf{u}$ into this then gives

$$a_h^{(0)}(\mathbf{u}, \mathbf{w}) = f_h(\mathbf{w}) + \sum_{F \in \mathcal{F}_h} \int_F \{\!\{ \nabla_h \mathbf{u} \}\!\} \cdot \mathbf{n}_F \cdot [\![\mathbf{w}]\!] \, \mathrm{d}\mathbf{s}.$$

In order to obtain consistency we modify the discrete bilinear form to

$$a_h^{(1)}(\mathbf{v}, \mathbf{w}) = \int_\Omega \nabla_h \mathbf{v} : \nabla_h \mathbf{w} \, \mathrm{d}\mathbf{x} - \sum_{F \in \mathcal{F}_h} \int_F \{\!\{ \nabla_h \mathbf{v} \}\!\} \cdot \mathbf{n}_F \cdot [\![\mathbf{w}]\!] \, \mathrm{d}\mathbf{s}.$$

Another property we would like to conserve from the continuous bilinear form is symmetry. In order to get a symmetric bilinear form we modify $a^{(1)}(\cdot,\cdot)$ to get

$$a_h^{sym}(\mathbf{v},\mathbf{w}) = \int_\Omega \nabla_h \mathbf{v} : \nabla_h \mathbf{w} \,\mathrm{d}\mathbf{x} - \sum_{F\in\mathcal{F}_h} \int_F \left(\{\!\{\nabla_h \mathbf{v}\}\!\} \cdot \mathbf{n}_F \cdot [\![\mathbf{w}]\!] + [\![\mathbf{v}]\!] \cdot \{\!\{\nabla_h \mathbf{w}\}\!\} \cdot \mathbf{n}_F\right) \mathrm{d}\mathbf{s}.$$

This is still a consistent bilinear form due to the continuity of the exact solution. Finally, for solvability of the discrete problem we need the discrete bilinear form to be coercive. Currently we have

$$a_h^{sym}(\mathbf{v},\mathbf{v}) = \|\nabla_h \mathbf{v}\|^2_{\mathcal{L}^2(\Omega)} - 2\sum_{F\in\mathcal{F}_h} \int_F \{\!\{\nabla_h \mathbf{v}\}\!\} \cdot \mathbf{n}_F \cdot [\![\mathbf{v}]\!] \,\mathrm{d}\mathbf{s}.$$

Since we do not have any a priori knowledge of the sign of the second term, we cannot expect discrete coercivity without the addition of further terms. In order to get discrete coercivity of the discrete bilinear form we penalise facets jumps, i.e.

$$\begin{aligned} a_h^{\mathrm{SIP}}(\mathbf{v},\mathbf{w}) = \int_\Omega \nabla_h \mathbf{v} : \nabla_h \mathbf{w} \,\mathrm{d}\mathbf{x} &- \sum_{F\in\mathcal{F}_h} \int_F \left(\{\!\{\nabla_h \mathbf{v}\}\!\} \cdot \mathbf{n}_F \cdot [\![\mathbf{w}]\!] + [\![\mathbf{v}]\!] \cdot \{\!\{\nabla_h \mathbf{w}\}\!\} \cdot \mathbf{n}_F\right) \mathrm{d}\mathbf{s} \\ &+ \sum_{F\in\mathcal{F}_h} \frac{\sigma}{h_F} \int_F [\![\mathbf{v}]\!] \cdot [\![\mathbf{w}]\!] \,\mathrm{d}\mathbf{s} \end{aligned} \tag{2.9}$$

for some parameter $\sigma > 0$ yet to be determined and the local facet diameter h_F. This bilinear form is still consistent in the sense of (2.8) since the exact solution is continuous.

Basic properties. The symmetric interior penalty bilinear form motivates the definition of the following discrete energy-norm

$$|\!|\!|\mathbf{v}|\!|\!|_e^2 := \|\nabla_h \mathbf{v}\|^2_{\mathcal{L}^2(\Omega)} + \sum_{F\in\mathcal{F}_h} \frac{\sigma}{h_F} \|[\![v]\!]\|^2_{\mathcal{L}^2(F)}$$

and the stronger norm

$$|\!|\!|\mathbf{v}|\!|\!|_{e,\sharp}^2 := |\!|\!|\mathbf{v}|\!|\!|_e^2 + \sum_{T\in\mathcal{T}_h} h_T \|\nabla \mathbf{v} \cdot \mathbf{n}_T\|^2_{\mathcal{L}^2(\partial T)}\,.$$

Lemma 2.13 (*Coercivity*). *Assume that the jump penalty parameter $\sigma > 0$ is sufficiently large. Then the bilinear form a_h^{SIP} is coercive on $\mathbf{V}_h$, i.e. there exists a constant $C_\sigma > 0$ such that*

$$a_h^{SIP}(\mathbf{v},\mathbf{v}) \geq C_\sigma |\!|\!|\mathbf{v}|\!|\!|_e^2 \qquad \textit{for all } \mathbf{v} \in \mathbf{V}_h.$$

Proof. See [Riv08, Lemma 6.6] or [DE11, Lemma 4.12]. □

Remark 2.14. Coercivity is given for $\sigma > C_{\mathrm{tr}}^2 N_\partial$, c.f. [DE11], where C_{tr} originates from the discrete inverse trace inequality and N_∂ is as defined by (2.1). On simplices, the constant C_{tr} is known to scale as $\sqrt{\frac{(k+1)(k+d)}{d}}$ [WH03], i.e. for coercivity we require $\sigma \sim k^2$.

Lemma 2.15 (*Boundedness*). *There exists a constant $C > 0$, independent of h, such that for all $\mathbf{v} \in \mathbf{V}^*$ and $\mathbf{w} \in \mathbf{V}_h$ it holds*

$$a_h^{SIP}(\mathbf{v},\mathbf{w}) \leq C |\!|\!|\mathbf{v}|\!|\!|_{e,\sharp} |\!|\!|\mathbf{w}|\!|\!|_e.$$

Proof. See [DE11, Lemma 4.16]. □

Non-Homogeneous Dirichlet Boundary Conditions. In the case of non-homogeneous Dirichlet boundary conditions

$$\mathbf{u} = \mathbf{g} \qquad \text{on } (0,T] \times \Gamma$$

on some part of the boundary $\Gamma \subseteq \partial\Omega$ we loose consistency of SIP bilinear form. Inserting the exact solution into (2.9) gives

$$a_h^{\mathrm{SIP}}(\mathbf{u}, \mathbf{v}_h) = -\int_\Gamma \mathbf{g}\nabla_h \mathbf{v}_h \cdot \mathbf{n}_F \,\mathrm{d}\mathbf{s} + \sum_{F\in\mathcal{F}_h^b} \frac{\sigma}{h_F} \int_{F\cap\Gamma} \mathbf{g}\cdot\mathbf{v}_h \,\mathrm{d}\mathbf{s}.$$

We can recover consistency by adding these terms to the right-hand side, i.e. we define a new right-hand side

$$f_h^D(\mathbf{v}_h) \coloneqq \int_\Omega \mathbf{f}\cdot\mathbf{v}_h \,\mathrm{d}\mathbf{x} - \int_\Gamma \mathbf{g}\nabla_h \mathbf{v}_h \cdot \mathbf{n}_F \,\mathrm{d}\mathbf{s} + \sum_{F\in\mathcal{F}_h^b} \frac{\sigma}{h_F} \int_{F\cap\Gamma} \mathbf{g}\cdot\mathbf{v}_h \,\mathrm{d}\mathbf{s}$$

such that $a_h^{\mathrm{SIP}}(\mathbf{u}, \mathbf{v}_h) = f_h^D(\mathbf{v}_h)$ holds for all $\mathbf{v}_h \in \mathbf{V}_h$.

2.2.2. Velocity-Pressure Coupling

Let $\mathcal{T}_h$ be an admissible decomposition of the domain Ω. The velocity pressure coupling appears in the momentum equation through the term $b(\mathbf{v}, q) = \int_\Omega (\nabla q)\cdot\mathbf{v}\,\mathrm{d}\mathbf{x}$. Now let $(\mathbf{v}, q) \in \boldsymbol{\mathcal{H}}_0^1(\Omega)\times\mathcal{H}^1(\Omega)$. To derive the discrete formulation we again begin by localising to each element and applying integration by parts

$$\begin{aligned} b_h^{(0)}(\mathbf{v}, q) &= \int_\Omega (\nabla_h q)\cdot\mathbf{v}\,\mathrm{d}\mathbf{x} \\ &= \sum_{T\in\mathcal{T}_h} \int_T (\nabla q)\cdot\mathbf{v}\,\mathrm{d}\mathbf{x} \\ &= -\sum_{T\in\mathcal{T}_h} \int_T (\nabla\cdot\mathbf{v}) q\,\mathrm{d}\mathbf{x} + \sum_{T\in\mathcal{T}_h} \int_{\partial T} \mathbf{v}\cdot\mathbf{n}_T q\,\mathrm{d}\mathbf{s}. \end{aligned}$$

As in the derivation of the SIP bilinear form we use that for all interior facets $F = \partial T_1 \cap \partial T_2$ we have $\mathbf{n}_F = \mathbf{n}_{T_1} = -\mathbf{n}_{T_2}$ to get

$$\begin{aligned} \sum_{T\in\mathcal{T}_h} \int_{\partial T} \mathbf{v}\cdot\mathbf{n}_T q\,\mathrm{d}\mathbf{s} &= \sum_{F\in\mathcal{F}_h^i} \int_F (\mathbf{v}|_{T_1}\cdot\mathbf{n}_{T_1}) q|_{T_1} + (\mathbf{v}|_{T_2}\cdot\mathbf{n}_{T_2}) q|_{T_2}\,\mathrm{d}\mathbf{s} + \sum_{F\in\mathcal{F}_h^b} \int_F \mathbf{v}\cdot\mathbf{n}_F q\,\mathrm{d}\mathbf{s} \\ &= \sum_{F\in\mathcal{F}_h^i} \int_F [\![\mathbf{v}\cdot\mathbf{n}_F q]\!]\,\mathrm{d}\mathbf{s} + \sum_{F\in\mathcal{F}_h^b} \int_F \mathbf{v}\cdot\mathbf{n}_F q\,\mathrm{d}\mathbf{s}. \end{aligned}$$

Together with our definition of boundary jumps this reduces to

$$\sum_{T\in\mathcal{T}_h} \int_{\partial T} \mathbf{v}\cdot\mathbf{n}_T q\,\mathrm{d}\mathbf{s} = \sum_{F\in\mathcal{F}_h} \int_F [\![\mathbf{v}\cdot\mathbf{n}_F q]\!]\,\mathrm{d}\mathbf{s}.$$

Following [Kan07], we make the design choice to replace q with the consistent flux $\hat{q} = \{\!\{q\}\!\}$ in the right-hand side of the above equation. This flux is chosen to make the problem symmetric, c.f. [Kan07, Section 4.1.3]. Due to our definition of boundary averages and jumps and since the average takes the same value on T_1 and T_2 we get the usual [DE11; Riv08] dG-form of the velocity-pressure coupling

$$b_h^{\mathrm{dG}}(\mathbf{v},q) = -\int_\Omega (\nabla_h \cdot \mathbf{v}) q \,\mathrm{d}\mathbf{x} + \sum_{F\in\mathcal{F}_h} \int_F [\![\mathbf{v}]\!] \cdot \mathbf{n}_F \{\!\{q\}\!\}.$$

For the case of $\boldsymbol{\mathcal{H}}(\mathrm{div})$-conforming FEM we can further simplify this by taking into account Lemma 2.10 together with the Dirichlet boundary conditions which gives us

$$b_h(\mathbf{v},q) = -\int_\Omega (\nabla_h \cdot \mathbf{v}) q \,\mathrm{d}\mathbf{x}$$

and since $\mathbf{v} \in \boldsymbol{\mathcal{H}}(\mathrm{div})$ also implies that $\nabla \cdot \mathbf{v} \in \mathcal{L}^2(\Omega)$ by Lemma 2.9, we can even take the continuous divergence operator and we get the standard velocity pressure coupling

$$b_h(\mathbf{v},q) = b(\mathbf{v},q) = -\int_\Omega (\nabla \cdot \mathbf{v}) q \,\mathrm{d}\mathbf{x}.$$

2.2.3. Upwinding for the Convection Part

Derivation. We begin by discretising the linearised convection term $c(\boldsymbol{\beta},\mathbf{v},\mathbf{w}) = \int_\Omega (\boldsymbol{\beta}\cdot\nabla)\mathbf{v}\cdot\mathbf{w}$ with the convective velocity field $\boldsymbol{\beta}$ for which we assume that $\nabla\cdot\boldsymbol{\beta} = 0$ holds pointwise and that $\boldsymbol{\beta}\cdot\mathbf{n} = 0$ holds on the boundary $\partial\Omega$. The idea is to preserve coercivity, i.e. we want that $c_h(\boldsymbol{\beta},\mathbf{v},\mathbf{v}) = 0$ holds. We again begin by localising to each element

$$c_h^{(0)}(\boldsymbol{\beta},\mathbf{v},\mathbf{w}) = \int_\Omega (\boldsymbol{\beta}\cdot\nabla_h)\mathbf{v}\cdot\mathbf{w}\,\mathrm{d}\mathbf{x} = \sum_{T\in\mathcal{T}_h} \int_T (\boldsymbol{\beta}\cdot\nabla)\mathbf{v}\cdot\mathbf{w}\,\mathrm{d}\mathbf{x}.$$

Using the divergence-free nature of the convective velocity and integration by parts, we observe that

$$\begin{aligned} 0 &= \int_T (\boldsymbol{\beta}\cdot\nabla)\mathbf{v}\cdot\mathbf{w}\,\mathrm{d}\mathbf{x} \\ &= -\int_T \boldsymbol{\beta}\cdot(\nabla(\mathbf{v}\cdot\mathbf{w}))\,\mathrm{d}\mathbf{x} - \int_{\partial T} (\boldsymbol{\beta}\cdot\mathbf{n}_T)\mathbf{v}\cdot\mathbf{w}\,\mathrm{d}\mathbf{s} \\ &= -\int_T \left[(\boldsymbol{\beta}\cdot\nabla)\mathbf{w}\cdot\mathbf{v} + (\boldsymbol{\beta}\cdot\nabla)\mathbf{v}\cdot\mathbf{w}\right]\mathrm{d}\mathbf{x} + \int_{\partial T} (\boldsymbol{\beta}\cdot\mathbf{n}_T)\mathbf{v}\cdot\mathbf{w}\,\mathrm{d}\mathbf{s}. \end{aligned} \tag{2.10}$$

In the final step we have used

$$\begin{aligned} \int_T \boldsymbol{\beta}\cdot(\nabla(\mathbf{v}\cdot\mathbf{w}))\,\mathrm{d}\mathbf{x} &= \int_T \boldsymbol{\beta}\cdot\left(\nabla\left(\sum_{i=1}^d v_i w_i\right)\right)\mathrm{d}\mathbf{x} \\ &= \int_T \sum_{j=1}^d \sum_{i=1}^d \left[\beta_j v_j \frac{\partial w_i}{\partial x_j} + \beta_j w_j \frac{\partial v_i}{\partial x_j}\right]\mathrm{d}\mathbf{x} \\ &= \int_T (\boldsymbol{\beta}\cdot\nabla)\mathbf{w}\cdot\mathbf{v}\,\mathrm{d}\mathbf{x} + \int_T (\boldsymbol{\beta}\cdot\nabla)\mathbf{v}\cdot\mathbf{w}\,\mathrm{d}\mathbf{x}. \end{aligned}$$

Rearranging (2.10) and testing with $\mathbf{w} = \mathbf{v}$ gives

$$\int_T (\boldsymbol{\beta}\cdot\nabla)\mathbf{v}\cdot\mathbf{v}\,\mathrm{d}\mathbf{x} = \frac{1}{2}\int_{\partial T} (\boldsymbol{\beta}\cdot\mathbf{n}_T)\mathbf{v}\cdot\mathbf{v}\,\mathrm{d}\mathbf{s}.$$

Therefore

$$
\begin{aligned}
c_h^{(0)}(\boldsymbol{\beta}, \mathbf{v}, \mathbf{v}) &= \sum_{T \in \mathcal{T}_h} \frac{1}{2} \int_{\partial T} (\boldsymbol{\beta} \cdot \mathbf{n}_T) \mathbf{v} \cdot \mathbf{v} \, \mathrm{d}s \\
&= \sum_{F \in \mathcal{F}_h^i} \frac{1}{2} \int_F (\boldsymbol{\beta} \cdot \mathbf{n}_F) [\![\mathbf{v} \cdot \mathbf{v}]\!] \, \mathrm{d}s + \sum_{F \in \mathcal{F}_h^b} \frac{1}{2} \int_F (\boldsymbol{\beta} \cdot \mathbf{n}_F) \mathbf{v} \cdot \mathbf{v} \, \mathrm{d}s.
\end{aligned}
$$

Lemma 2.4 gives $[\![\mathbf{v} \cdot \mathbf{v}]\!] = 2 [\![\mathbf{v}]\!] \cdot \{\!\{ \mathbf{v} \}\!\}$ and further using that $\boldsymbol{\beta} \cdot \mathbf{n}|_{\partial\Omega} = 0$ we have

$$
c_h^{(0)}(\boldsymbol{\beta}, \mathbf{v}, \mathbf{v}) = \sum_{F \in \mathcal{F}_h^i} \int_F (\boldsymbol{\beta} \cdot \mathbf{n}_F) [\![\mathbf{v}]\!] \cdot \{\!\{ \mathbf{v} \}\!\} \, \mathrm{d}s.
$$

So we get the discrete convective term

$$
c_h^{(1)}(\boldsymbol{\beta}, \mathbf{v}, \mathbf{w}) = \int_\Omega (\boldsymbol{\beta} \cdot \nabla_h) \mathbf{v} \cdot \mathbf{w} \, \mathrm{d}\mathbf{x} - \sum_{F \in \mathcal{F}_h^i} \int_F (\boldsymbol{\beta} \cdot \mathbf{n}_F) [\![\mathbf{v}]\!] \cdot \{\!\{ \mathbf{w} \}\!\} \, \mathrm{d}s.
$$

For further stability of this bilinear form, we penalise jumps of the discrete solution in a least squares sense. This can be interpreted as *upwinding* in terms of fluxes [DE11]. The new bilinear form is then given by

$$
\begin{aligned}
c_h(\boldsymbol{\beta}, \mathbf{v}, \mathbf{w}) = \int_\Omega (\boldsymbol{\beta} \cdot \nabla_h) \mathbf{v} \cdot \mathbf{w} \, \mathrm{d}\mathbf{x} &- \sum_{F \in \mathcal{F}_h^i} \int_F (\boldsymbol{\beta} \cdot \mathbf{n}_F) [\![\mathbf{v}]\!] \cdot \{\!\{ \mathbf{w} \}\!\} \, \mathrm{d}s \\
&+ \sum_{F \in \mathcal{F}_h^i} \int_F \frac{\gamma}{2} |\boldsymbol{\beta} \cdot \mathbf{n}_F| \, [\![\mathbf{v}]\!] \cdot [\![\mathbf{w}]\!] \, \mathrm{d}s
\end{aligned}
$$

where the (optional) stabilisation is controlled by the parameter $\gamma \geq 0$. In the case of $\gamma = 1$ this corresponds to upwinding.

Basic Properties. In connection with the convective term we introduce the norm for all $\mathbf{v} \in \mathbf{V}^*$

$$
|||\mathbf{v}|||_{\boldsymbol{\beta}}^2 = \|\mathbf{v}\|_{\mathcal{L}^2(\Omega)}^2 + |\mathbf{v}|_{\boldsymbol{\beta},\mathrm{upw}}^2
$$

with the upwind semi-norm

$$
|\mathbf{v}|_{\boldsymbol{\beta},\mathrm{upw}}^2 = \sum_{F \in \mathcal{F}_h^i} \int_F \frac{\gamma}{2} |\boldsymbol{\beta} \cdot \mathbf{n}_F| |[\![\mathbf{v}]\!]|^2 \, \mathrm{d}s
$$

which represents additional control over $\boldsymbol{\beta}$-scaled velocity jumps [SL17a]. Additionally, we define the stronger norm

$$
|||\mathbf{v}|||_{\boldsymbol{\beta},\sharp}^2 = |||\mathbf{v}|||_{\boldsymbol{\beta}}^2 + \|\boldsymbol{\beta}\|_\infty^2 \sum_{T \in \mathcal{T}_h} h_T^{-2} \|\mathbf{v}\|_{\mathcal{L}^2(T)}^2 + \|\boldsymbol{\beta}\|_\infty \sum_{T \in \mathcal{T}_h} \|\mathbf{v}\|_{\mathcal{L}^2(\partial T)}^2 .
$$

Lemma 2.16 (***Coercivity***)**.** *Let $\nabla \cdot \boldsymbol{\beta} = 0$ in Ω and $\boldsymbol{\beta} \cdot \mathbf{n} = 0$ on $\partial\Omega$. Then for all $\mathbf{v} \in \mathbf{V}_h$ we have that the convective term c_h is coercive with respect to $|\cdot|_{\boldsymbol{\beta},upw}$, i.e.*

$$
c_h(\boldsymbol{\beta}, \mathbf{v}, \mathbf{v}) = |\mathbf{v}|_{\boldsymbol{\beta},upw}^2.
$$

Proof. This follows by the construction of the convective term c_h. □

Lemma 2.17 (*Boundedness*). *Let $\varepsilon_1, \varepsilon_2 > 0$. There exists $C > 0$ such that for all $(\mathbf{v}, \mathbf{w}) \in \mathbf{V}^* \times \mathbf{V}_h$*

$$|c_h(\boldsymbol{\beta}, \mathbf{v}, \mathbf{w})| \leq C\left(\frac{1}{2\varepsilon_1} + \frac{1}{\varepsilon_2}\right) |||\mathbf{v}|||^2_{\boldsymbol{\beta},\sharp} + \frac{\varepsilon_1}{2}\|\mathbf{w}\|^2_{\mathcal{L}^2(\Omega)} + \varepsilon_2 |\mathbf{w}|^2_{\boldsymbol{\beta},upw}.$$

Proof. See [SL17a, Lemma 3.4]. □

2.2.4. The Spatially Semi-Discrete Problem

In order to formulate the continuous-in-time but spatially discretised problem, let us consider the following space-time spaces for the velocity and pressure

$$\mathbf{V}_h^T = \{\mathbf{v}_h \in \mathcal{L}^2(0, T; \mathbf{V}_h) \mid \partial_t \mathbf{v}_h \in \mathcal{L}^2(0, T; \mathbf{V}_h)\} \qquad \text{and} \qquad Q_h^T = \mathcal{L}^2(0, T; Q_h).$$

Then the spatially semi-discrete weak formulation of the Navier-Stokes equations is:

Problem P3 (*Spatially semi-discrete Navier-Stokes*). For $\mathbf{f} \in \mathcal{L}^2(0, T; \mathbf{V}')$ and $\mathbf{u}_{0h} = \mathbf{u}_h(0)$ find $(\mathbf{u}_h, p_h) \in \mathbf{V}_h^T \times Q_h^T$ such that for all $(\mathbf{v}_h, q_h) \in \mathbf{V}_h \times Q_h$ it holds

$$(\partial_t \mathbf{u}_h, \mathbf{v}_h)_{\mathcal{L}^2(\Omega)} + a_h^{\text{SIP}}(\mathbf{u}_h, \mathbf{v}_h) + c_h(\mathbf{u}_h, \mathbf{u}_h, \mathbf{v}_h) + b(\mathbf{v}_h, p_h) - b(\mathbf{u}_h, q_h) = \langle \mathbf{f}, \mathbf{v}_h \rangle_{\mathbf{V}', \mathbf{V}_h} \qquad (2.11)$$

where $\mathbf{u}_{0h}$ is an appropriate approximation of $\mathbf{u}_0$.

Lemma 2.18. *If the global spaces $\mathbf{V}_h$ and Q_h are* divergence conforming, *i.e.* (2.5) *holds, then the velocity approximation to Problem P3 is pointwise divergence free.*

Proof. Since $\nabla \cdot \mathbf{V}_h \subseteq Q_h$, we may test

$$0 = b_h(\mathbf{u}_h, q_h) = \int_\Omega q_h \nabla_h \cdot \mathbf{u}_h \, \mathrm{d}\mathbf{x}$$

with the test function $q_h = \nabla_h \cdot \mathbf{u}_h$. This gives

$$0 = b_h(\mathbf{u}_h, \nabla_h \cdot \mathbf{u}_h) = \int_\Omega (\nabla_h \cdot \mathbf{u}_h)^2 \, \mathrm{d}\mathbf{x} = \|\nabla_h \cdot \mathbf{u}_h\|^2_{\mathcal{L}^2(\Omega)}.$$

By the positive definiteness of norms, it follows that $\nabla_h \cdot \mathbf{u}_h = 0$ must hold. □

Remark 2.19. As a result of Assumption A2 and Lemma 2.18 the discrete velocity resulting from (2.11) is pointwise divergence free, i.e.

$$\mathbf{u}_h \in \{\mathbf{v} \in \mathbf{V}_h \mid \nabla \cdot \mathbf{v}_h(x) = 0, \forall \mathbf{x} \in \Omega\}.$$

Because of this and our assumption of homogeneous Dirichlet boundary conditions we may replace $\boldsymbol{\beta}$ with $\mathbf{u}_h$ and use the convective term $c_h(\cdot, \cdot, \cdot)$ as derived above.

Theorem 2.20. *Assume that the space pair $\mathbf{V}_h \times Q_h$ is inf-sup stable. Then Problem P3 has a unique solution $(\mathbf{u}_h, p_h) \in \mathbf{V}_h^T \times Q_h^T$. Furthermore, the discrete velocity admits the following energy estimate:*

$$\frac{1}{2}\|\mathbf{u}_h(T)\|^2_{\mathcal{L}^2(\Omega)} + \int_0^T C_\sigma |||\mathbf{u}_h(t)|||^2_e + |\mathbf{u}_h(t)|_{\mathbf{u}_h,upw} \, \mathrm{d}t \leq \|\mathbf{u}_{0h}\|^2_{\mathcal{L}^2} + \frac{3}{2}\|\mathbf{f}\|^2_{\mathcal{L}^1(0,T;\mathcal{L}^2)} \qquad (2.12)$$

with the discrete coercivity constant $C_\sigma > 0$ from Lemma 2.13.

Proof. We start by proving (2.12) following [DAL15; SL17a]. Testing (2.11) with $(\mathbf{u}_h, 0) \in \mathbf{V}_h^{\text{div}} \times Q_h$ gives

$$(\partial_t \mathbf{u}_h, \mathbf{u}_h)_{\mathcal{L}^2(\Omega)} + a_h^{\text{SIP}}(\mathbf{u}_h, \mathbf{u}_h) + c_h(\mathbf{u}_h, \mathbf{u}_h, \mathbf{u}_h) = (\mathbf{f}, \mathbf{v}_h)_{\mathcal{L}^2(\Omega)}.$$

Using the product rule we see that $\partial_t \|\mathbf{u}_h\|^2_{\mathcal{L}^2(\Omega)} = \partial_t (\mathbf{u}_h, \mathbf{u}_h)_{\mathcal{L}^2(\Omega)} = 2(\partial_t \mathbf{u}_h, \mathbf{u}_h)_{\mathcal{L}^2(\Omega)}$. Using this and the coercivity properties from Lemma 2.13 and Lemma 2.16 to estimate the left side from below and using the Cauchy-Schwarz inequality to estimate the right side from above, we get

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{u}_h\|^2_{\mathcal{L}^2(\Omega)} + C_\sigma |||\mathbf{u}_h|||_h^2 + |\mathbf{u}_h(t)|_{\mathbf{u}_h,\text{upw}} \leq \|\mathbf{f}\|_{\mathcal{L}^2(\Omega)} \|\mathbf{u}_h\|_{\mathcal{L}^2(\Omega)}. \tag{2.13}$$

Using the chain-rule and (2.13) we then see that

$$\|\mathbf{u}_h\|_{\mathcal{L}^2(\Omega)} \frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{u}_h\|_{\mathcal{L}^2(\Omega)} = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{u}_h\|^2_{\mathcal{L}^2(\Omega)} \leq \|\mathbf{f}\|_{\mathcal{L}^2(\Omega)} \|\mathbf{u}_h\|_{\mathcal{L}^2(\Omega)}$$

as a result of which we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{u}_h\|_{\mathcal{L}^2(\Omega)} \leq \|\mathbf{f}\|_{\mathcal{L}^2(\Omega)}. \tag{2.14}$$

Integrating (2.14) with respect to time then gives

$$\|\mathbf{u}_h(t)\|_{\mathcal{L}^2(\Omega)} \leq \|\mathbf{u}_{h0}\|_{\mathcal{L}^2(\Omega)} + \|\mathbf{f}\|_{\mathcal{L}^1(0,t;\mathcal{L}^2)} \leq \|\mathbf{u}_{h0}\|_{\mathcal{L}^2(\Omega)} + \|\mathbf{f}\|_{\mathcal{L}^1(0,T;\mathcal{L}^2)}.$$

Inserting this into the right-hand side of (2.13), integrating with respect to time and using Young's inequality then gives

$$\begin{aligned}\frac{1}{2}\left(\|\mathbf{u}_h(T)\|^2_{\mathcal{L}^2(\Omega)} - \|\mathbf{u}_{0h}\|^2_{\mathcal{L}^2(\Omega)}\right) + \int_0^T C_\sigma |||\mathbf{u}_h(t)|||_e^2 + |\mathbf{u}_h(t)|_{\mathbf{u}_h,\text{upw}} \,\mathrm{d}t \\ \leq \left(\|\mathbf{u}_{h0}\|_{\mathcal{L}^2(\Omega)} + \|\mathbf{f}\|_{\mathcal{L}^1(0,T;\mathcal{L}^2)}\right) \int_0^T \|\mathbf{f}\|_{\mathcal{L}^2(\Omega)} \,\mathrm{d}t \\ = \|\mathbf{u}_{h0}\|_{\mathcal{L}^2(\Omega)} \|\mathbf{f}\|_{\mathcal{L}^1(0,T;\mathcal{L}^2)} + \|\mathbf{f}\|^2_{\mathcal{L}^1(0,T;\mathcal{L}^2)} \\ \leq \frac{1}{2}\|\mathbf{u}_{h0}\|^2_{\mathcal{L}^2(\Omega)} + \frac{3}{2}\|\mathbf{f}\|^2_{\mathcal{L}^1(0,T;\mathcal{L}^2)}.\end{aligned}$$

Taking $1/2\,\|\mathbf{u}_{0h}\|^2_{\mathcal{L}^2(\Omega)}$ to the right-hand side we then obtain (2.12).

The ODE system (2.11) defining $\mathbf{u}_h$ is quadratic in the non-linearity and therefore (locally) Lipschitz [Joh16; Lay08]. Applying the theorem of Carathéodory (c.f. [Joh16, Theorem A.50]) gives us existence and uniqueness in some local time-interval $[0, t]$ with $t \leq T$. Due to the energy estimate (2.12) the solution cannot blow up for $t \in [0, T]$ which gives us the global statement [Joh16; Lay08].

The existence and uniqueness of the discrete pressure then follows from the discrete Babuška-Brezzi condition for the pair $\mathbf{V}_h \times Q_h$. For this, let

$$\mathbf{V}_h^{\text{div}} = \{\mathbf{v} \in \mathbf{V}_h \mid b(\mathbf{v}_h, q_h) = 0 \quad \forall q_h \in Q_h\}$$

then

$$(\mathbf{V}_h^{\text{div}})^\circ = \{\phi \in \mathbf{V}_h' \mid \phi(\mathbf{v}_h) = 0 \quad \forall \mathbf{v}_h \in \mathbf{V}_h^{\text{div}}\}.$$

We define

$$B' : Q_h \to (\mathbf{V}_h^{\text{div}})^\circ$$

by $B'q_h(\mathbf{v}_h) = b(\mathbf{v}_h, q_h)$. Now let $\mathbf{u}_h$ be the unique velocity solution. Then we define $\phi \in \mathbf{V}_h'$

$$\phi(\mathbf{v}_h) = (\mathbf{f}, \mathbf{v}_h)_{\mathcal{L}^2(\Omega)} - (\partial_t \mathbf{u}_h, \mathbf{v}_h)_{\mathcal{L}^2(\Omega)} - a_h^{\mathrm{SIP}}(\mathbf{u}_h, \mathbf{v}_h) - c_h(\mathbf{u}_h, \mathbf{u}_h, \mathbf{v}_h)$$

which is also in $(\mathbf{V}_h^{\mathrm{div}})^\circ$ since $\mathbf{u}_h$ is divergence free and so in $\mathbf{V}_h^{\mathrm{div}}$. Due to the inf-sup stability, we can apply Lemma 1.9 (iii) so there exists a unique $p_h \in Q_h$ such that

$$B'p_h = \phi.$$

□

Error estimates and convergence rate results. In order to obtain error estimates and convergence rates we additionally require the following assumptions (c.f. [SL17a; SLL$^+$18]).

Assumption A4. The global velocity space $\mathbf{V}_h$ has an optimal approximation property, i.e. there exits an approximation operator $\mathbf{j}_h : \mathbf{V} \to \mathbf{V}_h$ such that for all $\mathbf{v} \in \boldsymbol{\mathcal{H}}^r(\Omega)$ with $r > 2/3$ and $s = \min\{r, k+1\}$ it holds

$$\|\mathbf{v} - \mathbf{j}_h\mathbf{v}\|_{\mathcal{L}^2(T)} + h_T\|\mathbf{v} - \mathbf{j}_h\mathbf{v}\|_{\mathcal{H}^1(T)} \le Ch_T^s|\mathbf{v}|_{\mathcal{H}^s(T)}$$

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

Assumption A5. The approximation operator $\mathbf{j}_h$ fulfils the commuting diagram property:

$$\nabla \cdot (\mathbf{j}_h\mathbf{v}) = \pi_0 (\nabla \cdot \mathbf{v})$$

with the local orthogonal $\mathcal{L}^2$-projection $\pi_0 : \mathcal{L}^2(T) \to \mathbb{P}^k(T)$.

Assumption A6. The local space $\mathbf{V}^k(T)$ satisfies the discrete trace inequality, i.e, for all $\mathbf{v} \in \mathbf{V}^k(T)$ it holds

$$\|\mathbf{v}\|_{\mathcal{L}^2(\partial T)} \le C_{\mathrm{tr}} N_\partial^{1/2} h_T^{-1/2} \|\mathbf{v}\|_{\mathcal{L}^2(T)}, \qquad \forall T \in \mathcal{T}_h.$$

Assumption A7. We assume that Ω is a convex polygon for $d = 2$ or of class $\mathcal{C}^{1,1}$ for $d \in \{2,3\}$.

Theorem 2.21 (*Regularity of the Stokes problem*). *Let Assumption A7 hold true. Then for all $\mathbf{g} \in \boldsymbol{\mathcal{L}}^2(\Omega)$ the solution $(\mathbf{u}_s, p_s) \in \mathbf{V} \times Q$ of the stationary Stokes problem*

Find $(\mathbf{u}_s, p_s) \in \mathbf{V} \times Q$ s.t. $a(\mathbf{u}_s, \mathbf{v}) + b(\mathbf{v}, p_s) - b(\mathbf{u}_s, q) = (\mathbf{g}, \mathbf{v})_{\mathcal{L}^2}$ for all $(\mathbf{v}, q) \in \mathbf{V} \times Q$

also fulfils the regularity property $(\mathbf{u}_s, p_s) \in \boldsymbol{\mathcal{H}}^2 \times \mathcal{H}^1$ and the energy estimate $\sqrt{\nu}\|\mathbf{u}_s\|_{\mathcal{H}^2(\Omega)} + \|p\|_{\mathcal{H}^1(\Omega)} \le C\|\mathbf{g}\|_{\mathcal{L}^2(\Omega)}$.

Proof. See [BF13, Theorem IV.5.8]. □

Definition 2.22 (*Stationary Stokes projector*). Let $\mathbf{v} \in \boldsymbol{\mathcal{H}}^s(\mathcal{T}_h)$ for $r > 3/2$ such that $\nabla \cdot \mathbf{v} = 0$ holds pointwise. The *Stokes projection* $\pi_s\mathbf{v} \in \mathbf{V}_h^{\mathrm{div}}$ of $\mathbf{v}$ is the unique finite element solution of the problem

$$a_h(\pi_s\mathbf{v}, \mathbf{w}) = a(\mathbf{v}, \mathbf{w}) \qquad \forall\, \mathbf{w} \in \mathbf{V}_h^{\mathrm{div}}.$$

The projection operator π_s is called the *Stokes projector.*

Theorem 2.23 (*Stokes projection error estimate*). *Let Assumption A7 hold true. For $r > 3/2$ and $\mathbf{v} \in \boldsymbol{\mathcal{H}}^r$ we have with $s = \min\{r, k+1\}$ that there exists $C > 0$ such that*

$$\|\mathbf{v} - \pi_s \mathbf{v}\|_{\mathcal{L}^2(\Omega)} + h|||\mathbf{v} - \pi_s \mathbf{v}|||_{e,\sharp} \leq Ch \inf_{\mathbf{w}_h \in \mathbf{V}_h^{div}} |||\mathbf{v} - \mathbf{w}_h|||_{e,\sharp} \leq Ch^s |\mathbf{v}|_{\mathcal{H}^s(\Omega)}.$$

Proof. See [SL17a; SL17b]. □

Assumption A8. In the setting of Theorem 2.23 we assume in the $\boldsymbol{\mathcal{H}}^1$-conforming case that

$$\|\nabla_h \pi_s \mathbf{v}\|_{\mathcal{L}^\infty(\Omega)} \leq C \|\nabla_h \mathbf{v}\|_{\mathcal{L}^\infty(\Omega)}$$

and in the $\boldsymbol{\mathcal{H}}(\text{div})$-conforming case that

$$\|\mathbf{v} - \pi_s \mathbf{v}\|_{\mathcal{L}^\infty(\Omega)} + h\|\nabla_h \pi_s \mathbf{v}\|_{\mathcal{L}^\infty(\Omega)} \leq Ch \|\nabla_h \mathbf{v}\|_{\mathcal{L}^\infty(\Omega)}.$$

In Section 2.3 we will discuss the validity of these assumptions and present examples of spaces which fulfil them.

Theorem 2.24 (*Velocity discretisation error estimate*). *Let $\mathbf{u}$ be the solution to Problem P1 and $\mathbf{u}_h$ the solution of Problem P3. Let us additionally assume that $\mathbf{u} \in \mathcal{L}^2(0,T; \boldsymbol{\mathcal{H}}^r(\mathcal{T}_h)) \cap \mathcal{L}^1(0,T; \boldsymbol{\mathcal{W}}^{1,\infty}(\Omega))$ for some $r > 3/2$ and $\mathbf{u}_h(0) = \pi_s \mathbf{u}(0)$. With the error splitting*

$$\mathbf{u} - \mathbf{u}_h = (\mathbf{u} - \pi_s \mathbf{u}) - (\mathbf{u}_h - \pi_s \mathbf{u}_h) = \boldsymbol{\eta} - \mathbf{e}_h$$

we then have for the discretisation error $\mathbf{e}_h$ that

$$\frac{1}{2}\|\mathbf{e}_h\|^2_{\mathcal{L}^\infty(0,T;\mathcal{L}^2)} + \int_0^T \nu C_\sigma |||\mathbf{e}_h|||_e^2 + |\mathbf{e}_h|^2_{\mathbf{u}_h,upw} \,\mathrm{d}t$$
$$\leq e^{G_u(T)} \int_0^T \|\partial_t \boldsymbol{\eta}\|^2_{\mathcal{L}^2} + \|\mathbf{u}\|_{\mathcal{L}^\infty} \|\nabla_h \boldsymbol{\eta}\|^2_{\mathcal{L}^2} + (1 + Ch^{-2})\|\nabla \mathbf{u}\|_{\mathcal{L}^\infty} \|\boldsymbol{\eta}\|^2_{\mathcal{L}^2} \,\mathrm{d}t \quad (2.15)$$

with the Gronwall constant given by

$$G_u(T) = T + \|\mathbf{u}\|_{\mathcal{L}^1(0,T;\mathcal{L}^\infty(\Omega))} + C\|\nabla \mathbf{u}\|_{\mathcal{L}^1(0,T;\mathcal{L}^\infty)(\Omega)}. \quad (2.16)$$

Proof. See [SLL+18, Theorem 5.6]. □

Corollary 2.25 (*Velocity convergence rate*). *In the setting of Theorem 2.24, assume additionally that*

$$\mathbf{u} \in \{\mathbf{v} \in \mathcal{L}^2(0,T; \boldsymbol{\mathcal{H}}^\infty) \ : \ \partial_t \mathbf{v} \in \mathcal{L}^2(0,T; \boldsymbol{\mathcal{H}}^r)\}$$

for some $r > 3/2$. For $s = \min\{r, k+1\}$ and a constant $C > 0$ independent of ν and h we have

$$\frac{1}{2}\|\mathbf{e}_h\|^2_{\mathcal{L}^\infty(0,T;\mathcal{L}^2)} + \int_0^T \nu C_\sigma |||\mathbf{e}_h|||_e^2 + |\mathbf{e}_h|^2_{\mathbf{u}_h,upw} \,\mathrm{d}t$$
$$\leq Ch^{2(s-1)} e^{G_u(T)} \int_0^T h^2 |\partial_t \mathbf{u}|^2_{\mathcal{H}^s} + \Big[\|\mathbf{u}\|_{\mathcal{L}^\infty} + (h^2 + C)\|\nabla \mathbf{u}\|_{\mathcal{L}^\infty}\Big] |\mathbf{u}|^2_{\mathcal{H}^s} \,\mathrm{d}t. \quad (2.17)$$

Proof. See [SLL+18, Corollary 5.9]. □

Remark 2.26 (Pressure and Re-semi-robustness). We note that the error estimate in Theorem 2.24 is both pressure-robust and Re-semi-robust [SLL+18]. That is, the velocity error estimate is independent of the pressure and can therefore not be corrupted by poor pressure approximations (pressure-robust) and the Gronwall constant in (2.16) does not explicitly depend on the Reynolds number (Re-semi-robust).

2.3. Approximations of H(div)

Two examples of $\boldsymbol{\mathcal{H}}$(div)-conforming finite element spaces on simplicial meshes are the Brezzi-Douglas-Marini (BDM) spaces for $k \geq 1$ given by

$$\mathbb{BDM}_k := \{\mathbf{v}_h \in \boldsymbol{\mathcal{H}}_0(\mathrm{div}, \Omega) \mid \mathbf{v}_h|_T \in \mathbb{P}^k(T)\ \forall T \in \mathcal{T}_h\}$$

and the Raviart-Thomas (RT) spaces for $k \geq 0$ given by

$$\mathbb{RT}_k := \{\mathbf{v}_h \in \boldsymbol{\mathcal{H}}_0(\mathrm{div}, \Omega) \mid \mathbf{v}_h|_T \in \mathbb{RT}_k(T)\ \forall T \in \mathcal{T}_h\}$$

with the local Raviart-Thomas space

$$\mathbb{RT}_k(T) := \mathbb{P}^k(T) \oplus \boldsymbol{x}\mathbb{P}^k(T)$$

and $\boldsymbol{\mathcal{H}}_0(\mathrm{div}, \Omega) := \{\mathbf{v} \in \boldsymbol{\mathcal{H}}(\mathrm{div}; \Omega) \mid \mathbf{v} \cdot \mathbf{n}|_{\partial\Omega} = 0\}$. Pairing $\mathbf{V}_h = \mathbb{BDM}_k$ with $Q_h = \{q \in \mathcal{L}^2(\Omega) \mid q|_T \in \mathbb{P}^{k-1}(T)\ \forall T \in \mathcal{T}_h\}$ and pairing $\mathbf{V}_h = \mathbb{RT}_k$ with $Q_h = \{q \in \mathcal{L}^2(\Omega) \mid q|_T \in \mathbb{P}^k(T)\ \forall T \in \mathcal{T}_h\}$ are known to give inf-sup stable pairs, i.e.

$$\inf_{q_h \in Q_h} \sup_{\mathbf{v}_h \in \mathbf{V}_h} \frac{\int_\Omega (\nabla \cdot \mathbf{v}_h) q_h \,\mathrm{d}\mathbf{x}}{\|\mathbf{v}_h\|_{\boldsymbol{\mathcal{H}}(\mathrm{div};\Omega)} \|q_h\|_{\mathcal{L}^2(\Omega)}} \geq \beta_h > \beta_0 > 0$$

holds with $\beta_0 > 0$ independent of h [JLM$^+$17] and independent of k [LS17]. From the definition of these spaces we can conclude that these pairs are also divergence conforming, i.e.

$$\nabla \cdot \mathbf{V}_h \subseteq Q_h.$$

Furthermore, these choices of velocity approximation spaces satisfy the assumptions made in Chapter 2. For further details see for example [BBF13, Proposition 2.5.4] or [JLM$^+$17] for Assumption A4, see [BBF13, Proposition 2.5.2] for Assumption A5 and [DE11] for Assumption A6. The validity of Assumption A8 is still an open problem in the $\boldsymbol{\mathcal{H}}$(div)-conforming context, c.f. [SLL$^+$18, Remark 5.4].

Remark 2.27. On rectangular or rectangular hexahedral meshes $\mathcal{K}_h$ consider the local space $\mathbb{P}_{k_1,k_2}(K) := \{p(x_1, x_2) \mid p(x_1, x_2) = \sum_{i \leq k_1, j \leq k_2} x_1^i x_2^j\}$. Then the Raviart-Thomas space is

$$\mathbf{V}_h = \mathbb{RT}_{[k]} := \{\mathbf{v}_h \in \boldsymbol{\mathcal{H}}_0(\mathrm{div}, \Omega) \mid \mathbf{v}_h|_K \in \mathbb{RT}_{[k]}(K)\ \forall K \in \mathcal{K}_h\}$$

where the local Raviar-Thomas space on rectangles and rectangular hexahedrons is defined by

$$\mathbb{RT}_{[k]}(K) = \begin{cases} \mathbb{P}_{k+1,k}(K) \times \mathbb{P}_{k,k+1}(K) & d = 2 \\ \mathbb{P}_{k+1,k,k}(K) \times \mathbb{P}_{k,k+1,k}(K) \times \mathbb{P}_{k,k,k+1}(K) & d = 3. \end{cases}$$

Together with the pressure space

$$Q_h = \mathbb{Q}^k_{disc} = \begin{cases} \mathbb{P}_{k,k}(K) & d = 2 \\ \mathbb{P}_{k,k,k}(K) & d = 3 \end{cases}$$

this pair also builds an inf-sup stable finite element pair which is divergence conforming satisfying our assumptions, c.f. [BBF13; CKS06; SST02]. For the case $d = 3$ the space $\mathbb{P}_{k_1,k_2,k_3}(K)$ is defined analogous to the two dimensional case.

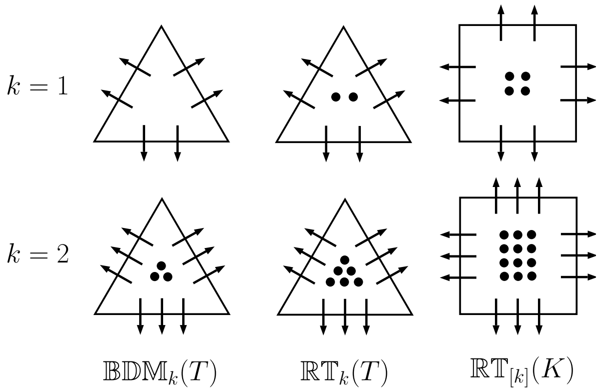


Figure 2.3.: BDM space on triangles and RT space on triangles and rectangles.

Remark 2.28. The dimension of the local $\mathbb{BDM}_k(T)$ spaces is given by

$$\dim \mathbb{BDM}_k(T) = \begin{cases} (k+1)(k+2) & d=2 \\ \frac{1}{2}(k+1)(k+2)(k+3) & d=3 \end{cases}$$

on simplices and the dimension of the local $\mathbb{RT}_k(T)$ spaces is given by

$$\dim \mathbb{RT}_k(T) = \begin{cases} (k+1)(k+3) & d=2 \\ \frac{1}{2}(k+1)(k+2)(k+4) & d=3 \end{cases}$$

on simplices. On rectangular or rectangular hexahedral elements the dimension of the local Raviart-Thomas space is

$$\dim \mathbb{RT}_{[k]}(K) = \begin{cases} 2(k+1)(k+2) & d=2 \\ 3(k+1)^2(k+2) & d=3. \end{cases}$$

The degrees of freedom for the low order cases $k \in \{0,1,2\}$ can bee seen in Figure 2.3. For further details on the construction of these spaces we refer to [BBF13].

Remark 2.29. Another approach for $\boldsymbol{\mathcal{H}}(\mathrm{div})$-conforming FEM on simplicial meshes has been to enrich BDM and RT spaces locally on each element with divergence-free polynomials. Further enrichment has been based on using rational functions. Here the tangential components of the basis functions across interfaces are modified to ensure tangential continuity. For more details, see [BBF13, Sec. 8.9.1.1].

Furthermore, [CKS05] use a completely discontinuous approach and reconstruct an exactly divergence free velocity in $\boldsymbol{\mathcal{H}}(\mathrm{div};\Omega)$ via a post-processing step.

Remark 2.30. In the $\boldsymbol{\mathcal{H}}^1$-conforming case we have the Scott-Vogelius pair $\mathbb{P}^k/\mathbb{P}^{k-1}_{disc}$ on barycenter refined simplicial meshes is for $k \geq d$ as an inf-sup stable FE pair [Qin94; Zha04] which is divergence conforming.

3. IMEX Multistep Schemes

The spatial discretisation of the time-dependent Navier-Stokes equations (1.1) as described in Chapter 2 leads to a large initial value DAE system of the form

$$M\dot{\mathbf{u}}_h = A(\mathbf{u}_h, p_h) + C(\mathbf{u}_h)$$

where A is the linear Stokes part and right-hand side, C is the non-linear convective term in the Navier-Stokes equations and M is the mass matrix, i.e.

$$\begin{aligned} Mu_h &= (u_h, v_h)_{\mathcal{L}^2(\Omega)} \\ A(\mathbf{u}_h, p_h) &= -a_h^{\mathrm{SIP}}(\mathbf{u}_h, \mathbf{v}_h) - b_h(\mathbf{v}_h, p_h) + b_h(\mathbf{u}_h, q_h) + \mathbf{f}_h(\mathbf{v}_h) \\ C(\mathbf{u}_h) &= -c_h(\mathbf{u}_h, \mathbf{u}_h, \mathbf{v}_h) \end{aligned}$$

in $\mathbf{V}_h \times Q_h$. For ease of notation, we will drop the explicit dependence on h here.

Discretising the time derivative by some finite difference scheme then results in a time-stepping scheme. In order to avoid the necessity of solving a non-linear system of equations in each time step we want to treat the convective term explicitly. On the other hand, the diffusion part is very stiff and should therefore be treated implicitly so as to avoid very small time steps. Furthermore, the implicit treatment of the divergence constraint is essential for us as it ensures the pointwise divergence feee nature of the spatial discretisation considered in Chapter 2. As a result of this we want to treat the Stokes part implicitly. The implicit treatment of A and explicit treatment of C leads to so-called IMEX schemes [ARW95].

In this chapter we will derive IMEX multistep methods, i.e. time stepping methods which consider multiple previous steps to advance the solution. We will consider schemes up to formal order three. We will then analyse these schemes by considering a scalar test problem as is customary in ODE analysis, and extend the analysis from the literature to schemes formal order three. We note that there is no direct implication from the scalar test problem and our large system. However, this analysis gives us some heuristics about the relative restrictive nature of the schemes.

3.1. Derivation

We follow [ARW95] to derive general multistep IMEX schemes of order $s \geq 1$. For a fixed time step $\tau > 0$, let us therefore consider the general s-step linear multistep IMEX scheme

$$\frac{1}{\tau}\sum_{j=-1}^{s-1} a_j M\mathbf{u}^{n-j} = \sum_{j=-1}^{s-1} b_j A(\mathbf{u}^{n-j}) + \sum_{j=0}^{s-1} c_j C(\mathbf{u}^{n-j}). \tag{3.1}$$

for an differential equation of the form

$$M\dot{\mathbf{u}} = A(\mathbf{u}) + C(\mathbf{u}) \tag{3.2}$$

where C is the part which we want to treat explicitly and A the part we want to treat implicitly. Note that we have "hidden" the pressure p in an additional component of $\mathbf{u}$ in order to make the derivation more concise. The operators M, A and C are therefore extended to act on the appropriate components of $\mathbf{u}$. Note that M is therefore not invertible, i.e. we have a DAE system.

Given that $\mathbf{u}$, A and C are sufficiently smooth we can build a Taylor expansion around $t_n = \tau n$. For this we we observe that

$$\begin{aligned}\mathbf{u}(t_n+\tau) &= \mathbf{u}(t_n) + \tau\dot{\mathbf{u}}(t_n) + \frac{1}{2}\tau^2\ddot{\mathbf{u}}(t_n) + \dots\\ \mathbf{u}(t_n) &= \mathbf{u}(t_n)\end{aligned}$$

and

$$\mathbf{u}(t_n-k\tau) = \mathbf{u}(t_n) - k\tau\dot{\mathbf{u}}(t_n) + \frac{1}{2}(k\tau)^2\ddot{\mathbf{u}}(t_n) - \dots .$$

With this a Taylor expansion of (3.1) around $t_n = \tau n$ gives

$$\begin{aligned}\frac{1}{\tau}\left[\sum_{j=-1}^{s-1} a_j\right] M\mathbf{u}(t_n) + \left[a_{-1} - \sum_{j=1}^{s-1} j a_j\right] M\dot{\mathbf{u}}(t_n) + \dots + \frac{\tau^{p-1}}{p!}\left[a_{-1} + \sum_{j=1}^{s-1}(-j)^p a_j\right] M\mathbf{u}^{(p)}(t_n)\\ - \sum_{j=0}^{s-1} c_j C(\mathbf{u}(t_n)) + \tau\sum_{j=1}^{s-1} j c_j \left.\frac{\mathrm{d}C}{\mathrm{d}t}\right|_{t=t_n} - \dots - \frac{\tau^{p-1}}{(p-1)!}\sum_{j=1}^{s-1}(-j)^{p-1}c_j \left.\frac{\mathrm{d}^{p-1}C}{\mathrm{d}t^{p-1}}\right|_{t=t_n}\\ - \sum_{j=-1}^{s-1} b_j A(\mathbf{u}(t_n)) - \tau\left[b_{-1} - \sum_{j=1}^{s-1} j b_j\right]\left.\frac{\mathrm{d}A}{\mathrm{d}t}\right|_{t=t_n} - \dots\\ - \frac{\tau^{p-1}}{(p-1)!}\left[b_{-1} + \sum_{j=1}^{s-1}(-j)^{p-1} b_j\right]\left.\frac{\mathrm{d}^{p-1}A}{\mathrm{d}t^{p-1}}\right|_{t=t_n} + \mathcal{O}(\tau^p).\end{aligned}$$

Inserting (3.2) into this gives us a scheme of order p provided that

$$\begin{aligned}\sum_{j=-1}^{s-1} a_j &= 0\\ a_{-1} - \sum_{j=1}^{s-1} j a_j &= \sum_{j=0}^{s-1} c_j = \sum_{j=-1}^{s-1} b_j = 1\\ \frac{1}{2}\left[a_{-1} + \sum_{j=1}^{s-1} j^2 a_j\right] &= -\sum_{j=1}^{s-1} j c_j = b_{-1} - \sum_{j=1}^{s-1} j b_j \qquad (3.3)\\ &\vdots\\ \frac{1}{p!}\left[a_{-1} + \sum_{j=1}^{s-1}(-j)^p a_j\right] &= \sum_{j=1}^{s-1}\frac{(-j)^{p-1}c_j}{(p-1)!} = \frac{1}{(p-1)!}\left[b_{-1} + \sum_{j=1}^{s-1}(-j)^{p-1} b_j\right].\end{aligned}$$

Theorem 3.1. *For the s-step IMEX scheme* (3.1) *it holds that*

(a) *The* $2p+2$ *constraints of the system* (3.3) *are linearly independent if* $p \le s$. *There therefore exist IMEX schemes of order* s.

(b) An s-step IMEX scheme has maximal order of accuracy equal to s.

(c) The family of s-step IMEX schemes of order s has s parameters.

Proof. This proof is based on [Ruu93] where the constraint $\sum_{j=-1}^{s-1} b_j = 1$ was lost due to the additional assumption $a_{-1} = 1$.

To show (a) and (c) we only need to consider the case $p = s$ since the linear independence for $p < s$ is then clear. For $\mathbf{x} = [a_{-1}, a_0, \dots, s_{s-1}, b_{-1}, b_0, \dots, b_{s-1}, c_0, c_1, \dots, c_{s-1}]^T \in \mathbb{R}^{3s+2}$ and $\mathbf{b} = [1, 0, 0 \dots, 0]^T \in \mathbb{R}^{2s+2}$ we can rewrite the system (3.3) as $A\mathbf{x} = \mathbf{b}$ with A given as

$$A = \begin{bmatrix}
0 & \cdots & 0 & 1 & 1 & \cdots & 1 & \\
 & & & 0 & 0 & \cdots & 0 & \\
 & & & -1 & & & & \\
 & V_1^T & & -2 & & -D_sV_2^T & & 0 \\
 & & & \vdots & & & & \\
 & & & -s & & & & \\
 & & & 1 & & & & \\
 & 0 & & \vdots & & V_2^T & & \text{-}V_2^T \\
 & & & 1 & & & &
\end{bmatrix}$$

where

$$V_1 = \begin{bmatrix}
1 & 1 & 1 & \cdots & 1 \\
1 & 0 & 0 & \cdots & 0 \\
1 & -1 & 1 & \cdots & (-1)^s \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & 1-s & (1-s)^2 & \cdots & (1-s)^s
\end{bmatrix} \quad \text{and} \quad V_2 = \begin{bmatrix}
1 & 0 & \cdots & 0 \\
1 & -1 & \cdots & (-1)^{s-1} \\
\vdots & \vdots & \ddots & \vdots \\
1 & 1-s & \cdots & (1-s)^{s-1}
\end{bmatrix}$$

are the Vandermonde matrices of the sequences $\{x_j^1 = 2 - j\}_{j=1}^{s+1}$ and $\{x_j^2 = 1 - j\}_{j=1}^{s}$ whose elements are each pairwise different. The matrix D_s is the diagonal matrix

$$D_s = \begin{bmatrix}
1 & 0 & \cdots & 0 \\
0 & 2 & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & s
\end{bmatrix}.$$

The two Vandermonde matrices have non-vanishing determinant [Fis14] and have therefore full rank while D_s is also of full rank. From this we see that the rows and columns of A are linearly independent. Therefore, the $2p + 2$ constraints of the system (3.3) are linearly independent and the system admits a $(3s + 2) - (2p + 2) \overset{p=s}{=} s$ parameter family of solutions.

To show (b) let us assume that we have an s-step IMEX scheme of order $s + r$ for some

$r \geq 1$. From the Taylor expansion above, we see that for such a scheme it must hold that

$$\begin{aligned}\sum_{j=0}^{s-1} c_j &= \sum_{j=-1}^{s-1} b_j \\ \sum_{j=1}^{s-1}(-j)c_j &= b_1 + \sum_{j=1}^{s-1}(-j)b_j \\ &\vdots \\ \sum_{j=1}^{s-1}(-j)^s c_j &= b_1 + \sum_{j=1}^{s-1}(-j)^s b_j.\end{aligned}$$

By setting $d_j = c_j - b_j$ we can rewrite this as

$$\begin{aligned}b_1 + \sum_{j=0}^{s-1} d_j &= 0 \\ b_1 + \sum_{j=1}^{s-1}(-j)d_j &= 0 \\ &\vdots \\ b_1 + \sum_{j=1}^{s-1}(j)^s d_j &= 0.\end{aligned} \tag{3.4}$$

As before we can rewrite (3.4) in matrix form as $B\mathbf{y} = \mathbf{0}$ with $\mathbf{y} = [b_{-1}, d_0, d_1, \ldots, d_{s-1}]^T$ and

$$B = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 0 & -1 & \cdots & 1-s \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & (-1)^s & \cdots & (1-s)^s \end{bmatrix}.$$

We observe that $B = V_1^T$ which is non-singular. Therefore we have $\ker(B) = \{\mathbf{0}\}$ which in turn implies that $b_{-1} = 0$ and $c_j = b_j$ for $j = 0, 1, \ldots s-1$. However, for the scheme to be IMEX it is necessary for $b_{-1} \neq 0$. Therefore a s-step IMEX scheme cannot have order greater than s. □

3.1.1. IMEX Schemes of Order 1

Setting $p = s = 1$, (3.3) gives the restrictions

$$\begin{aligned} a_{-1} + a_0 = 0 \\ 1 = a_{-1} = c_0 = b_{-1} + b_0. \end{aligned}$$

This gives the one-parameter family of first-order IMEX schemes

$$\frac{1}{\tau}[M\mathbf{u}^{n+1} - M\mathbf{u}^n] = [(1-\gamma)A(\mathbf{u}^n) + \gamma A(\mathbf{u}^{n+1})] + C(\mathbf{u}^n) \tag{3.5}$$

where we restrict $0 \leq \gamma \leq 1$ to prevent a large truncation error.

The choice $\gamma = 0$ leads to the explicit Euler scheme

$$M\mathbf{u}^{n+1} - M\mathbf{u}^n = \tau A(\mathbf{u}^n) + \tau C(\mathbf{u}^n)$$

which is fully explicit and which we will not consider here.

SBDF1 Scheme. Setting $\gamma = 1$ gives the scheme

$$M\mathbf{u}^{n+1} - M\mathbf{u}^n = \tau A(\mathbf{u}^{n+1}) + \tau C(\mathbf{u}^n) \tag{3.6}$$

which applies the implicit Euler scheme to A and the explicit Euler scheme to C. This scheme has been considered for example in [MT98, Section 18]. Schemes such as this where the time derivative together with the implicit part are treated with a *backward differentiation formula* (BDF), c.f. [HV03], and some extrapolation formula is applied to the explicit part, are called semi-implicit BDF (SBDF) schemes. The scheme (3.6) will therefore be referred to as *SBDF1*.

3.1.2. IMEX Schemes of Order 2

Setting $p = s = 2$ in (3.3) gives the restrictions

$$\begin{aligned} a_{-1} + a_0 + a_1 &= 0 \\ 1 = a_{-1} - a_1 &= c_0 + c_1 = b_{-1} + b_0 + b_1 \\ \frac{1}{2}[a_{-1} + a_1] &= -c_1 = b_{-1} - b_1 \end{aligned}$$

which leaves two free parameters. We choose $-c_1 = \gamma$. It then follows that $c_0 = \gamma + 1$, $a_{-1} = \gamma + 1/2$, $a_1 = \gamma - 1/2$ and $a_0 = -2\gamma$. By further choosing $b_1 = \delta$, we get $b_{-1} = \gamma + \delta$ and $b_0 = 1 - \gamma - 2\delta$. This gives the family of second order IMEX scheme

$$\begin{aligned} \frac{1}{\tau}\left[\left(\gamma + \frac{1}{2}\right) M\mathbf{u}^{n+1} - 2\gamma M\mathbf{u}^n + \left(\gamma - \frac{1}{2}\right) M\mathbf{u}^{n-1}\right] &= (\gamma + 1)C(\mathbf{u}^n) - \gamma C(\mathbf{u}^{n-1}) \\ &+ (\gamma + \delta)A(\mathbf{u}^{n+1}) + (1 - \gamma - 2\delta)A(\mathbf{u}^n) + \delta A(\mathbf{u}^{n-1}). \end{aligned} \tag{3.7}$$

We will consider the following choices for (γ, δ) which give the following schemes:

SBDF2 Scheme. Setting $(\gamma, \delta) = (1, 0)$ in (3.7) gives the splitting method

$$\frac{1}{2\tau}[3M\mathbf{u}^{n+1} - 4M\mathbf{u}^n + M\mathbf{u}^{n-1}] = A(\mathbf{u}^{n+1}) + 2C(\mathbf{u}^n) - C(\mathbf{u}^{n-1}).$$

This is again a SBDF scheme and we will therefore refer to it as *SBDF2*.

CNAB Scheme. Setting $(\gamma, \delta) = (1/2, 0)$ gives

$$\frac{1}{\tau}[M\mathbf{u}^{n+1} - M\mathbf{u}^n] = \frac{1}{2}\left[A(\mathbf{u}^{n+1}) + A(\mathbf{u}^n)\right] + \frac{3}{2}C(\mathbf{u}^n) - \frac{1}{2}C(\mathbf{u}^{n-1})$$

which applies the Crank-Nicolson method to the implicit part and the second-order Adams-Bashforth method to the explicit part. It is thus called the *CNAB* (Crank-Nicolson, Adams-Bashforth) splitting method. This method has also been considered, e.g. in [MT98, Section 19]

		Coefficients							
Scheme	(γ, δ)	a_{-1}	a_0	a_1	b_{-1}	b_0	b_1	c_0	c_1
SBDF2	$(1, 0)$	$3/2$	-2	$1/2$	1	0	0	2	-1
CNAB	$(1/2, 0)$	1	-1	0	$1/2$	$1/2$	0	$3/2$	$-1/2$
CNAB(1/16)	$(1/2, 1/16)$	1	-1	0	$9/16$	$3/8$	$1/16$	$3/2$	$-1/2$
CNAB(1/4)	$(1/2, 1/4)$	1	-1	0	$3/4$	0	$1/4$	$3/2$	$-1/2$
CNLF	$(0, 1/2)$	$1/2$	0	$-1/2$	$1/2$	0	$1/2$	1	0

Table 3.1.: Second order IMEX Coefficients.

CNAB(δ) Schemes. The choice $\gamma = 1/2$ which leads to the second order Adams-Bashforth method to the explicit part has been considered in the cited literature together with other choices of δ. Due to the similarity of these schemes to the CNAB scheme, we shall refer to them as *CNAB(δ)*. The choice $\delta = 1/16$ has been considered by [ARW95; FHV97] which gives the scheme

$$\frac{1}{\tau}[M\mathbf{u}^{n+1} - M\mathbf{u}^n] = \frac{9}{16}A(\mathbf{u}^{n+1}) + \frac{3}{8}A(\mathbf{u}^n) + \frac{1}{16}A(\mathbf{u}^{n-1}) + \frac{3}{2}C(\mathbf{u}^n) - \frac{1}{2}C(\mathbf{u}^{n-1})$$

and [FHV97; NL79] considered the choice $\delta = 1/4$ which in turn gives the scheme

$$\frac{1}{\tau}[M\mathbf{u}^{n+1} - M\mathbf{u}^n] = \frac{3}{4}A(\mathbf{u}^{n+1}) + \frac{1}{4}A(\mathbf{u}^{n-1}) + \frac{3}{2}C(\mathbf{u}^n) - \frac{1}{2}C(\mathbf{u}^{n-1}).$$

CNLF Scheme. The choice $(\gamma, \delta) = (0, 1/2)$ in (3.7) gives the scheme

$$\frac{1}{2\tau}[M\mathbf{u}^{n+1} - M\mathbf{u}^{n-1}] = \frac{1}{2}\left[A(\mathbf{u}^{n+1}) + A(\mathbf{u}^{n-1})\right] + C(\mathbf{u}^n)$$

which [ARW95] refer to as *CNLF* (Crank-Nicolson, Leap Frog) since it applies the Leap Frog scheme to the time derivative together with the explicit part and a scheme similar to Crank-Nicolson to the implicit part.

To summarise, the coefficients of the second order IMEX multistep schemes are given in Table 3.1.

3.1.3. IMEX Schemes of Order 3

Setting $p = s = 3$ in (3.3) gives the set of restrictions

$$\begin{gathered}
a_{-1} + a_0 + a_1 + a_2 = 0 \\
1 = a_{-1} - a_1 - 2a_2 = c_0 + c_1 + c_2 = b_{-1} + b_1 + b_1 + b_2 \\
\frac{1}{2}[a_{-1} + a_1 + 4a_2] = -c_1 - 2c_2 = b_{-1} - b_1 - 2b_2 \\
\frac{1}{6}[a_{-1} - a_1 - 8a_2] = \frac{1}{2}[c_1 + 4c_2] = \frac{1}{2}[b_{-1} + b_1 + 4b_2].
\end{gathered}$$

The parametrisation (γ, δ, η) chosen by [ARW95] gives the following family of schemes:

$$\begin{aligned}\frac{1}{\tau}\Bigg[\left(\frac{1}{2}\gamma^2+\gamma+\frac{1}{3}+\delta\right)M\mathbf{u}^{n+1}+\left(\frac{3}{2}\gamma^2-2\gamma+\frac{1}{2}-\delta\right)M\mathbf{u}^{n}+\left(\frac{3}{2}\gamma^2+\gamma-1\right)M\mathbf{u}^{n-1}\\ +\left(-\frac{1}{2}\gamma^2+\frac{1}{6}\right)M\mathbf{u}^{n-2}\Bigg]=\left(\frac{\gamma^2+3\gamma}{2}+1+\frac{23}{12}\delta\right)C(u^n)-\left(\gamma^2+2\gamma+\frac{4}{3}\delta\right)C(u^{n-1})\\ +\left(\frac{\gamma^2+\gamma}{2}+\frac{5}{12}\delta\right)C(u^{n-2})+\left(\frac{\gamma^2+\gamma}{2}+\eta\right)A(\mathbf{u}^{n+1})+\left(1-\gamma^2-3\eta+\frac{23}{12}\delta\right)A(\mathbf{u}^{n})\\ +\left(\frac{\gamma^2-\gamma}{2}+3\eta-\frac{4}{3}\delta\right)A(\mathbf{u}^{n-1})+\left(\frac{5}{12}\delta-\eta\right)A(\mathbf{u}^{n-2}).\end{aligned}$$

SBDF3 Scheme. The choice $(\gamma, \delta, \eta) = (1, 0, 0)$ gives the scheme:

$$\frac{1}{\tau}\left[\frac{11}{6}M\mathbf{u}^{n+1}-3M\mathbf{u}^{n}+\frac{3}{2}M\mathbf{u}^{n-1}-\frac{1}{3}M\mathbf{u}^{n-2}\right]=A(\mathbf{u}^{n+1})+3C(\mathbf{u}^n)-3C(u^{n-1})+C(\mathbf{u}^{n-2})$$

which applies the 3rd order BDF formula to the implicit term and an extrapolation to the explicit term. We will therefore refer to it as the *SBDF3* scheme.

3.2. Analysis: Restrictions on the Explicit Eigenvalue

For the analysis of the methods derived in Section 3.1 we will follow the approach taken by [FHV97]. We will therefore consider the scalar test equation

$$\dot{x} = \alpha x + \beta x \tag{3.8}$$

with $\alpha, \beta \in \mathbb{C}$ to analyse the stability properties of the schemes. This differs to the more restrictive approach taken by [ARW95] where the test equation $\dot{x} = \alpha x + i\beta x$ with $\alpha, \beta \in \mathbb{R}$ was used. In our context α and β represent the eigenvalues of implicit part A and the explicit part C respectively [FHV97]. Note that this implies that both operators A and C are *linear*. This means that this section is only directly applicable to the Oseen problem which can be seen as a linearised variant of the Navier-Stokes problem and appears as an auxiliary problem to the Navier-Stokes equations [Joh16, Chapter 5]. Applying the general IMEX scheme (3.1) to the scalar test problem (3.8) gives rise to the linear difference equation

$$\sum_{j=-1}^{s-1} a_j x^{n-j} = \tau \sum_{j=-1}^{s-1} b_j \alpha x^{n-j} + \tau \sum_{j=0}^{s-1} c_j \beta x^{n-j}. \tag{3.9}$$

Let $\zeta \in \mathbb{C}$. Inserting $x^{n-j} = (\zeta)^{s-j-1}$ into (3.9) and making the substitutions $\tau\alpha \to \lambda$ and $\tau\beta \to \mu$ gives us the *characteristic equation*

$$\Phi(\zeta) := \sum_{j=-1}^{s-1} a_j \zeta^{s-1-j} - \lambda \sum_{j=-1}^{s-1} b_j \zeta^{s-1-j} - \mu \sum_{j=0}^{s-1} c_j \zeta^{s-1-j}. \tag{3.10}$$

It is known, c.f. [HNW93, Section III.3], that the solution of the linear difference equation (3.9) has the form

$$x^{n+1} = p_1(n)(\zeta_1)^n + p_2(n)(\zeta_2)^n + \cdots + p_l(n)(\zeta_l)^n$$

where $\zeta_1, \dots, \zeta_l$ are the roots of $\Phi(\zeta)$ with respective multiplicity $m_1, \dots, m_l$ and $p_j(n)$ are polynomials of degree $m_j - 1$.

To obtain boundedness for x^{n+1} as $n \to \infty$ we require that $|\zeta_j| \leq 1$ and for $|\zeta_j| = 1$ we additionally need that $m_j = 1$. This motivates the following definitions (see for example [HNW93; HW96]):

Definition 3.2. The multistep method (3.1) is called *stable* if the roots of the characteristic polynomial (3.10) satisfy the following conditions:

1. All roots of $\Phi(\zeta)$ lie in or on the unit circle in $\mathbb{C}$;
2. All roots on the unit circle are simple.

Now dividing $\Phi(\zeta)$ by ζ^s and substituting $z = {}^1\!/\!{}_\zeta$ the characteristic equation becomes

$$\Psi(z) := \sum_{j=-1}^{s-1} a_j z^{j+1} - \lambda \sum_{j=-1}^{s-1} b_j z^{j+1} - \mu \sum_{j=0}^{s-1} c_j z^{j+1}. \tag{3.11}$$

For ease of notation we denote

$$E(z) = \sum_{j=-1}^{s-1} a_j z^{j+1}, \quad F(z) = \sum_{j=-1}^{s-1} b_j z^{j+1} \quad \text{and} \quad G(z) = \sum_{j=0}^{s-1} c_j z^{j+1}$$

so that the modified characteristic equation becomes $\Psi(z) = E(z) - \lambda F(z) - \mu G(z)$. For stability it is therefore necessary for the roots of (3.11) to satisfy $|z| \geq 1$ with strict inequality for multiple roots. The stability function must therefore not vanish for all $z < 1$, i.e. a necessary condition for stability is

$$\Psi(z) \neq 0 \quad \text{for all} \quad |z| < 1. \tag{3.12}$$

This condition is also sufficient if we omit the necessity for roots with $|z| = 1$ to be single. As in [FHV97] we therefore take (3.12) as a criterium for stability and check separately whether multiple roots with modulus 1 occur.

Definition 3.3. The set

$$\mathcal{S} = \{\mu \in \mathbb{C} \mid \text{The explicit method (i.e. } \lambda = 0\text{) is stable.}\}$$

is called the stability region of the explicit part of the IMEX scheme.

We can characterise the boundary $\partial\mathcal{S}$ of the stability region of the explicit method by setting $\lambda = 0$ in (3.11), inserting $z = e^{i\theta}$ and rearranging for μ to get the so called *root locus curve* [HW96]:

$$\mu^{\mathcal{S}}(\theta) = \frac{E(e^{i\theta})}{G(e^{i\theta})} \quad \text{for } \theta \in [\pi, \pi].$$

Then the interior $\mathrm{int}(\mathcal{S})$ of the region of stability $\mathcal{S}$ where all roots have modulus strictly less than 1, is given by the complement of the set $\{\mu(z) \ : \ |z| \leq 1\}$.

Definition 3.4. Let $c_j = 0$ for $j = 0, \dots, s-1$. The remaining implicit multistep scheme is called *A-stable* if for the set

$$\tilde{\mathcal{S}} = \{\lambda \in \mathbb{C} \mid \text{The (implicit) method is stable.}\}$$

it holds that $\tilde{\mathcal{S}} \subset \mathbb{C}^- = \{z \in \mathbb{C} \mid \ \mathrm{Re}(z) \leq 0\}$.

Definition 3.5. The set

$$\mathcal{D} = \Big\{\mu \in \mathbb{C} \mid \text{The IMEX scheme is stable for all } \lambda \in \mathbb{C}^-.\Big\}$$

is the stability region of the explicit part of the IMEX scheme for which A-stability is preserved for the implicit part of the scheme.

We observe $\mathcal{D} \subset \overline{\mathcal{S}}$. To characterise the boundary of $\mathcal{D}$ we use the following lemma:

Lemma 3.6. *Let* $M(\theta) := {}^{E(e^{i\theta})}/_{F(e^{i\theta})}$ *and* $N(\theta) := {}^{G(e^{i\theta})}/_{F(e^{i\theta})}$. *Suppose that* $\operatorname{Re}(N(\theta)) \not\equiv 0$ *and that* $M(\theta)$, $N(\theta)$ *are bounded for all* $\theta \in [-\pi, \pi]$. *Then* $\partial\mathcal{D} \subset \{\mu^{\mathcal{D}}(\theta) \,:\, \theta \in [-\pi, \pi]\}$ *with*

$$\mu^{\mathcal{D}}(\theta) := \frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{M(\theta) + M(-\theta)}{N(-\theta)}\right)\left[\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{N(\theta)}{N(-\theta)}\right)\right]^{-1}.$$

Proof. See [FHV97, Lemma 3.2]. □

As a result of Lemma 3.6 we have a description of the boundary $\partial\mathcal{D}$ under the assumption that $F(e^{i\theta})$ does not vanish and $\operatorname{Re}(N(\theta)) \not\equiv 0$ for $\theta \in [-\pi, \pi]$. This means that we can plot both $\partial\mathcal{S}$ and $\partial\mathcal{D}$ in order to get a sense of how much of a restriction, the demand for A-stability of the implicit part, is on the explicit part of the individual IMEX scheme.

3.2.1. 1st Order Schemes

SBDF1 Scheme. For the SBDF1 and CNLF schemes we can show that we have A-stability for the implicit eigenvalue (i.e. $\lambda \in \mathbb{C}^-$) if the explicit eigenvalue is in the stability domain $\mathcal{S}$ of the explicit scheme. To this end, let

$$\varphi_\mu(z) := \frac{E(z) - \mu G(z)}{F(z)}.$$

Then the stability criterium (3.12) becomes

$$\lambda \neq \varphi_\mu(z) \quad \text{for all} \quad |z| < 1. \tag{3.13}$$

If we look at the SBDF1 scheme we have the coefficients $a_{-1} = 1,\ a_0 = -1,\ b_{-1} = 1,\ b_0 = 0$ and $c_0 = 1$. This gives

$$E(z) = 1 - z, \quad F(z) = 1 \quad \text{and} \quad G(z) = z$$

For stability of the explicit method, i.e. the explicit Euler method ($\lambda = 0$), we require for the root of the modified characteristic equation

$$\Psi(z) = 1 - z - \mu z$$

to have modulus greater or equal than 1. Setting $\Psi(z) = 0$ we get that $z = {}^1/_{1+\mu}$ which gives the stability domain of the explicit Euler scheme

$$\mathcal{S} = \{\mu \in \mathbb{C} \,:\, |1 + \mu| \leq 1\}.$$

If we then look at

$$\varphi_\mu(z) = 1 - z - \mu z$$

we see that

$$\operatorname{Re}(\varphi_\mu(z)) = \operatorname{Re}(1 - z(1 + \mu)) > 0$$

for all $|z| < 1$ and $\mu \in \mathcal{S}$, giving A-stability for the implicit eigenvalues so long as the explicit eigenvalues are within the stability domain of the explicit Euler scheme.

3.2.2. 2nd Order Schemes

SBDF2 Scheme. For the SBDF2 scheme we have $a_{-1} = 3/2,\ a_0 = -2,\ a_1 = -1/2,\ b_{-1} = 1,\ b_0 = 0,\ b_1 = 0,\ c_0 = 2$ and $c_1 = -1$ which gives us

$$E(z) = \frac{1}{2}(z-3)(z-1), \quad F(z) = 1 \quad \text{and} \quad G(z) = z(2-z).$$

This in turn yields the root locus curve for the stability domain of the explicit method

$$\mu^{\mathcal{S}}(\theta) = \frac{(e^{i\theta}-3)(e^{i\theta}-1)}{2e^{i\theta}(2-e^{i\theta})}.$$

To compute the stability domain of the explicit method in order to obtain A-stability of the implicit method we have

$$M(\theta) = \frac{1}{2}(e^{i\theta}-3)(e^{i\theta}-1) \quad \text{and} \quad N(\theta) = e^{i\theta}(2-e^{i\theta}).$$

Applying Lemma 3.6 gives us the root locus curve, characterised by

$$\begin{aligned}
\mu^{\mathcal{D}}(\theta) &= \frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{M(\theta)+M(-\theta)}{N(-\theta)}\right)\left[\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{N(\theta)}{N(-\theta)}\right)\right]^{-1} \\
&= \frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{(e^{i\theta}-3)(e^{i\theta}-1)+(e^{-i\theta}-3)(e^{-i\theta}-1)}{2e^{-i\theta}(2-e^{-i\theta})}\right)\left[\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{e^{i\theta}(2-e^{i\theta})}{e^{-i\theta}(2-e^{-i\theta})}\right)\right]^{-1} \\
&= \frac{ie^{i\theta}(e^{i\theta}-1)^3(3e^{i\theta}-1)}{(1-2e^{i\theta})^2}\left[\frac{-6ie^{3i\theta}(e^{i\theta}-1)^2}{(1-2e^{i\theta})^2}\right]^{-1} \\
&= -\frac{1}{6}e^{-2i\theta}(e^{i\theta}-1)(3e^{i\theta}-1) \\
&= -\frac{1}{6}(1-e^{-i\theta})(3-e^{-i\theta})
\end{aligned}$$

for $\theta \in [\pi, -\pi]$ where the derivatives have been computed using Wolfram|Alpha [Wol18a; Wol18b]. In Figure 3.1 we see these stability regions in the complex plane. Here we observe, that in this case, the restriction for A-stability is not much stronger than just requiring stability in the explicit part.

CNAB(δ) Schemes. The Crank-Nicolson/Adams-Bashforth type schemes have the coefficients $a_{-1} = 1,\ a_0 = -1,\ a_1 = 0,\ b_{-1} = 1/2 + \delta,\ b_0 = 1/2 - 2\delta,\ b_1 = \delta,\ c_0 = 3/2$ and $c_1 = -1/2$. The resulting polynomials are given by

$$E(z) = 1-z, \quad F(z) = \delta z^2 + \frac{1}{2}(1-4\delta)z + \frac{1}{2} + \delta \quad \text{and} \quad G(z) = \frac{1}{2}z(3-z).$$

This gives the locus root curve of the explicit two-step Adams-Bashforth method

$$\mu^{\mathcal{S}}(\theta) = \frac{2(1-e^{i\theta})}{e^{i\theta}(3-e^{i\theta})}$$

for $\theta \in [-\pi, \pi]$.

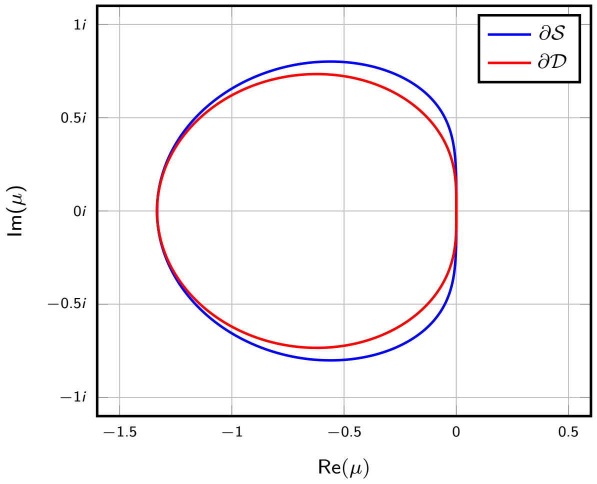


Figure 3.1.: Stability Region of the SBDF2 Scheme.

The choice $\delta = \frac{1}{16}$ yields the polynomials

$$M(\theta) = \frac{16(1-e^{i\theta})}{e^{2i\theta}+6e^{i\theta}+9} \quad \text{and} \quad N(\theta) = \frac{8e^{i\theta}(3-e^{i\theta})}{e^{2i\theta}+6e^{i\theta}+9}.$$

Applying Lemma 3.6 to parametrise the boundary of the stability domain $\mathcal{D}$ we get

$$\begin{aligned}
\mu^{\mathcal{D}}(\theta) &= \frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{M(\theta)+M(-\theta)}{N(-\theta)}\right)\left[\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{N(\theta)}{N(-\theta)}\right)\right]^{-1} \\
&= \frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{\frac{16(1-e^{i\theta})}{e^{2i\theta}+6e^{i\theta}+9}+\frac{16(1-e^{-i\theta})}{e^{-2i\theta}+6e^{-i\theta}+9}}{\frac{8e^{-i\theta}(3-e^{-i\theta})}{e^{-2i\theta}+6e^{-i\theta}+9}}\right)\left[\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{\frac{8e^{i\theta}(3-e^{i\theta})}{e^{2i\theta}+6e^{i\theta}+9}}{\frac{8e^{-i\theta}(3-e^{-i\theta})}{e^{-2i\theta}+6e^{-i\theta}+9}}\right)\right]^{-1} \\
&= \frac{2ie^{i\theta}(e^{i\theta}-1)^3(3e^{2i\theta}+34e^{i\theta}-5)}{(1-3e^{i\theta})^2(3+e^{i\theta})^3}\,\frac{-(3+e^{i\theta})^3(1-3e^{i\theta})^2}{9ie^{i\theta}(e^{i\theta}-1)^2(1+3e^{i\theta})(e^{2i\theta}+10e^{i\theta}+1)} \\
&= -\frac{2(e^{i\theta}-1)(3e^{2i\theta}+34e^{i\theta}-5)}{9(1+3e^{i\theta})(e^{2i\theta}+10e^{i\theta}+1)}
\end{aligned}$$

where the derivatives have been determined using Wolfram|Alpha [Wol18c; Wol18d].

The choice $\delta = \frac{1}{4}$ yields the polynomials

$$M(\theta) = \frac{4(1-e^{i\theta})}{e^{2i\theta}+3} \quad \text{and} \quad N(\theta) = \frac{2e^{i\theta}(3-e^{i\theta})}{e^{2i\theta}+3}.$$

This in turn gives the locus root curve parametrised by

$$
\begin{aligned}
\mu^{\mathcal{D}}(\theta) &= \frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{M(\theta)+M(-\theta)}{N(-\theta)}\right)\left[\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{N(\theta)}{N(-\theta)}\right)\right]^{-1} \\
&= \frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{\frac{4(1-e^{i\theta})}{e^{2i\theta}+3}+\frac{4(1-e^{-i\theta})}{e^{-2i\theta}+3}}{\frac{2e^{-i\theta}(3-e^{-i\theta})}{e^{-2i\theta}+3}}\right)\left[\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{\frac{2e^{i\theta}(3-e^{i\theta})}{e^{2i\theta}+3}}{\frac{2e^{-i\theta}(3-e^{-i\theta})}{e^{-2i\theta}+3}}\right)\right]^{-1} \\
&= \frac{2ie^{i\theta}(e^{i\theta}-1)^3(3e^{3i\theta}+7e^{2i\theta}+25e^{i\theta}-3)}{(1-3e^{i\theta})^2(3+e^{2i\theta})^2}\frac{-(1-3e^{i\theta})^2(3+e^{2i\theta})^2}{3ie^{i\theta}(e^{i\theta}-1)^2(3e^{4i\theta}+4e^{3i\theta}+34e^{2i\theta}+4e^{i\theta}+3)} \\
&= -\frac{2(e^{i\theta}-1)(3e^{3i\theta}+7e^{2i\theta}+25e^{i\theta}-3)}{3(3e^{4i\theta}+4e^{3i\theta}+34e^{2i\theta}+4e^{i\theta}+3)}
\end{aligned}
$$

where the derivates have again been determined using Wolfram|Alpha [Wol18e; Wol18f]. The three stability regions for the explicit part of the CNAB(δ) schemes can be seen in Figure 3.2. Here we see that in the case of $\delta = 1/4$, requiring A-stability is not significantly stronger a condition than simply requiring stability of the explicit part. However, for the case $\delta = 1/16$ we see that $\mathcal{D}$ is significantly smaller than $\mathcal{S}$.

To compute the explicit stability domain for A-stability of the implicit part with $\delta = 0$ we are unable to utilise Lemma 3.6 since $M(\theta) = 2(1-e^{i\theta})/(1+e^{i\theta})$ is unbounded for θ close to π and $-\pi$. However, we have

$$
\begin{align}
\varphi_\mu(z) &= \frac{2(1-z)-\mu z(3-z)}{1+z} \tag{3.14} \\
&= \frac{2+2z-z-z^2-3z+z^2-\mu z(3-z)}{1+z} \notag \\
&= \frac{(2-z)(1+z)-(1+\mu)z(3-z)}{1+z} \notag \\
&= 2-z-(1+\mu)\frac{z(z-3)}{1+z}. \tag{3.15}
\end{align}
$$

Let us define $\chi(z) := -z(z-3)/1+z$. We then observe

$$
\begin{aligned}
\mathrm{Re}\left(\chi(e^{i\theta})\right) &= \frac{\cos^3\theta - 2\cos^2\theta + \sin^2\theta\cos\theta - 3\cos\theta - 4\sin^2\theta}{\sin^2\theta + (\cos\theta+1)^2} \\
&= \frac{\cos\theta\sin^2\theta + \cos^3\theta + 2\cos^2\theta + \cos\theta - 2\sin^2\theta - 2\cos^2\theta - 4\cos\theta - 2(\sin^2\theta+\cos^2\theta)}{\sin^2\theta + (\cos\theta+1)^2} \\
&= \frac{(cos^{\theta}-2)(\sin^2\theta + (\cos\theta+1)^2)}{\sin^2\theta + (\cos\theta+1)^2} \\
&= \cos\theta - 2.
\end{aligned}
$$

For $\theta \to \pi$ we then have $\mathrm{Re}(\chi(e^{i\theta})) \to -3$, but $\left|\chi(e^{i\theta})\right| \to \infty$. Therefore χ maps the unit circle to the half plane $\{\xi \in \mathbb{C} \ : \ \mathrm{Re}(\xi) \geq -3\}$ and the imaginary axis lies entirely in this image. This means that if $\mathrm{Im}(1+\lambda) \neq 0$, the image of the unit disk under φ_μ and $\mathbb{C}^-$ have a non-empty intersection which in turn means that $\lambda \in \mathbb{R}$ for λ to be in $\mathcal{D}$. Now for real λ

$$
\begin{aligned}
\mathrm{Re}(\varphi_\mu(e^{i\theta)}) &= 2 - \cos\theta + (1+\lambda)(\cos\theta - 2) \\
&= -\lambda(2-\cos\theta)
\end{aligned}
$$

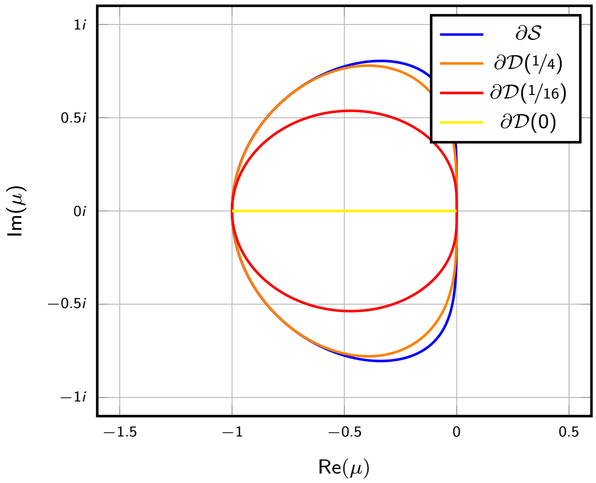


Figure 3.2.: Stability Region of the CNAB(δ) Schemes.

so $\mathrm{Re}(\varphi_\mu(e^{i\theta})) \geq 0$ if $\lambda \leq 0$. Furthermore, for $\lambda \leq 0$ we see from (3.15) that the unit circle is mapped into $\mathbb{C}^-$ if additionally $1 + \lambda \geq 0$. This gives the domain $\mathcal{D} = [-1, 0]$ and $\mathcal{D}$ is indeed a subset of $\mathcal{S}$.

CNLF Scheme. For the CNLF scheme we can proceed similarly as for the SBDF1 scheme. Here $a_{-1} = 1/2,\ a_0 = 0,\ a_1 = -1/2,\ b_{-1} = 1/2,\ b_0 = 0,\ b_1 = 1/2,\ c_0 = 1$ and $c_1 = 0$. This gives

$$E(z) = \frac{1}{2}(1 - z^2), \quad F(z) = \frac{1}{2}(1 + z^2) \quad \text{and} \quad G(z) = z$$

which in turn yields the root locus curve

$$\begin{aligned}\mu^{\mathcal{S}}(\theta) &= \frac{1 - e^{2i\theta}}{2e^{i\theta}}\\ &= \frac{1}{2}(e^{-i\theta} - e^{i\theta})\\ &= -i\sin(\theta)\end{aligned}$$

for $\theta \in [-\pi, \pi]$. So μ has to be restricted to $[-i, i]$, however, we have to check whether multiple roots occur on the boundary. Setting $\mu = i$ we get the modified characteristic equation for the explicit method

$$\begin{aligned}\Psi(z) &= \frac{1}{2}(1 - z^2) - iz\\ &= -\frac{1}{2}(z + 1)^2\end{aligned}$$

which has the double root $z_{1,2} = -i$. Furthermore, $\mu = -1$ gives

$$\begin{aligned}\Psi(z) &= \frac{1}{2}(1 - z^2) + iz\\ &= -\frac{1}{2}(z - 1)^2\end{aligned}$$

which has the double root $z_{1,2} = i$. Therefore, the region of stability for the explicit part of the method is

$$\mathcal{S} = (-i, i).$$

Let us now look at $\varphi_\mu(z)$ together with $\mu = -i\sin(\theta)$:

$$\begin{aligned}\varphi_\mu(z) &= \frac{1 - z^2 + 2iz\sin(\theta)}{1 + z^2} \\ &= \frac{1 - z^2 + 2iz\sin(\theta)}{1 + z^2}\frac{1 + \bar{z}^2}{1 + \bar{z}^2} \\ &= \frac{1 - (z^2 - \bar{z}^2) - |z|^4 + 2iz\sin(\theta)(1 + \bar{z}^2)}{1 + z^2 + \bar{z}^2 + |z|^4}.\end{aligned}$$

The denominator is clearly real and positive for $|z| < 1$. Now with $z = a+ib$ we can characterise the numerator by

$$\begin{aligned}\mathrm{Re}(1 - (z^2 - \bar{z}^2) + 2iz\sin(\theta)(1 + \bar{z}^2)) &= 1 + |z|^4 - 2\sin(\theta)\,\mathrm{Im}(z(1 + \bar{z}^2)) \\ &= 1 + |z|^4 - 2\sin\theta\,\mathrm{Im}((a + ib)(1 + a^2 - 2iab - b^2)) \\ &= 1 + |z|^4 - 2\sin\theta(-2a^2b + a^2b + b - b^3) \\ &= 1 + |z|^4 - 2\sin\theta(b(1 - |z|^2)) \\ &= (1 - |z|^2)(1 + |z|^2 - 2b\sin(\theta))\end{aligned}$$

and this is strictly positive for $|z| < 1$. So A-stability for the implicit eigenvalue is given provided that the explicit eigenvalue is in $\mathcal{S}$.

Remark 3.7. The result $\mathcal{S} = (-i, i)$ shows that this scheme is inappropriate for our numerical studies since our choice of discrete convective term $c_h(\boldsymbol{\beta}, \mathbf{u}, \mathbf{v})$ is only skew-symmetric in the $\boldsymbol{\mathcal{H}}^1$-conforming context where both the jump term and upwinding term vanish. In the $\boldsymbol{\mathcal{H}}(\mathrm{div})$-conforming contex, these terms do not vanish and due to our choice of using upwind fluxes the corresponding operator has eigenvalues with a non-vanishing real part. The CNLF scheme is therefore inherently unstable in our context and will therefore not be considered in the numerical experiments.

3.2.3. 3rd Order Schemes

SBDF3. The coefficients of the SBDF3 scheme are given by $a_{-1} = {}^{11}\!/_{16}$, $a_0 = -3$, $a_1 = {}^3\!/_2$, $a_2 = {}^{-1}\!/_3$, $b_{-1} = 1$, $b_0 = 0$, $b_1 = 0$, $b_2 = 0$, $c_0 = 3$, $c_1 = -3$ and $c_2 = 1$. This gives us gives us

$$E(z) = \frac{11}{6} - 3z + \frac{3}{2}z^2 - \frac{1}{3}z^3, \quad F(z) = 1 \quad \text{and} \quad G(z) = 3z - 3z^2 + z^3.$$

The locus root curve of the explicit part of the SBDF3 scheme is given by

$$\mu^{\mathcal{S}}(\theta) = \frac{11 - 18e^{i\theta} + 9e^{2i\theta} - 2e^{3i\theta}}{6(3e^{i\theta} - 3e^{2i\theta} + e^{3i\theta})}$$

for $\theta \in [-\pi, \pi]$. The plot of this in the complex plane can be seen in Figure 3.3. Due to the second Dahlquist barrier, c.f. [HW96, Theorem V.1.6], any A-stable multistep scheme is at most of order two, so the BDF3 scheme is not A-stable. Therefore there is no restriction $\mathcal{D}$ within $\mathcal{S}$ for which we get A-stability for the implicit part of the SBDF3 scheme.

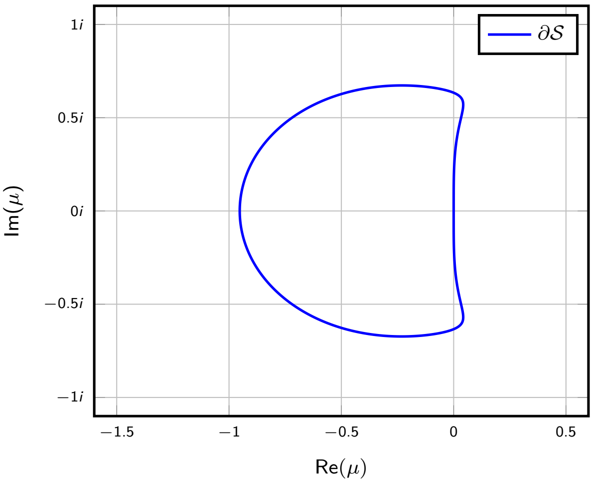


Figure 3.3.: Stability Region of the SBDF3 Scheme.

3.3. Analysis: Restrictions on the Implicit Eigenvalue

Requiring full A-stability is a severe restriction on the methods considered. We will therefore consider the less demanding stability restriction of $A(\alpha)$-stability since this is sufficient in most practical situations [FHV97].

Definition 3.8. A linear multistep method is said to be $A(\alpha)$-*stable* if for $0 < \alpha < \pi/2$ the set

$$\mathcal{W}_\alpha = \{\zeta \in \mathbb{C} \, : \, |\arg(-\zeta)| < \alpha\}$$

is a subset of the stability region of the method [HW96].

We will now consider for which values of α we have stability for arbitrary $\mu \in \mathcal{S}$ given that $\lambda \in \mathcal{W}_\alpha$.

Lemma 3.9. *Let E, F and G be the parts of the modified characteristic polynomial* (3.11) *corresponding to the time-derivative, implicit and explicit part of an IMEX scheme respectively. Suppose we have*

$$|\arg(E(z) - \mu G(z))| \leq \frac{\pi}{2} + \beta \quad \text{and} \quad |\arg(F(z))| \leq \gamma$$

for all $|z| = 1$ and $\mu \in \partial\mathcal{S}$, with $\beta + \gamma < \pi/2$. Then the IMEX scheme is stable for any $\mu \in \mathcal{S}$ and $\lambda \in \mathcal{W}_\alpha$, with $\alpha = \pi/2 - \beta - \gamma$.

Proof. See [FHV97, Lemma 4.1]. □

3.3.1. 2nd Order Schemes

To determine the angle β in Lemma 3.9 we observe that for schemes of order two, we have

$$E(z) - \mu G(z) = a_{-1}(1 - \rho_1 z)(1 - \rho_2 z) \tag{3.16}$$

where ρ_1 and ρ_2 are the roots of the characteristic polynomial of the explicit part of the method. For $\mu \in \partial\mathcal{S}$ we then have $|\rho_1| = 1$ and $|\rho_2| \leq r$ for some constant $r \leq 1$ determined by the explicit method. "It follows by geometrical considerations that we can take $\beta = \arcsin(r)$" [FHV97].

SBDF2 Scheme. For the SBDF2 scheme, (3.16) reads

$$\frac{1}{2}(z-3)(z-1)-\mu z(2-z)=\frac{3}{2}(1-\rho_1 z)(1-\rho_2 z).$$

If $\mu \in \partial\mathcal{S}$ we may write $\rho_1 = e^{i\theta}$. Inserting this into the above equation then yields

$$(\frac{1}{2}+\mu)z^2-2(\mu+1)z+\frac{3}{2}=\frac{3}{2}(\rho_2 e^{i\theta}z^2-(\rho_2+e^{i\theta})z+1)$$

and since $\partial\mathcal{S}$ is parametrised by $\mu(\theta) = {}^{(e^{i\theta}-3)(e^{i\theta}-1)}/_{2e^{i\theta}(2-e^{i\theta})}$, reparametrisation $\theta \mapsto -\theta$ and comparing coefficients yields

$$\rho_2 = \frac{2-3e^{i\theta}}{3-6e^{i\theta}}.$$

So $|\rho_2| \leq {}^5/_9$. Furthermore, $\arg(F(z)) = \arg(1) = 0$ so we can apply Lemma 3.9 and get the lower bound for the stability angle

$$\alpha = \frac{\pi}{2} - \arcsin(\frac{5}{9}) \approx 0.31\pi.$$

Note that this is significantly smaller than ${}^\pi/_2$, the value for α of the BDF2 multistep method [HW96].

CNAB(δ) Schemes. For this set of schemes (3.16) reads

$$(1-2)-\frac{\mu}{2}z(3-z)=(1-\rho_1 z)(1-\rho_2 z).$$

For $\mu \in \partial\mathcal{S}$ we have $\rho_1 = e^{i\theta}$ which yields

$$\frac{\mu}{2}z^2-(1+\frac{3\mu}{2})z+1=\rho_2 e^{i\theta}z^2-(\rho_1+\rho_2)z+1.$$

For $\mu \in \partial\mathcal{S}$ we also know that $\mu(\theta) = {}^{2(1-e^{i\theta})}/_{e^{i\theta}(3-e^{i\theta})}$. Then reparametrising $\mu(\theta)$ with $\theta \mapsto -\theta$ and comparing coefficients gives

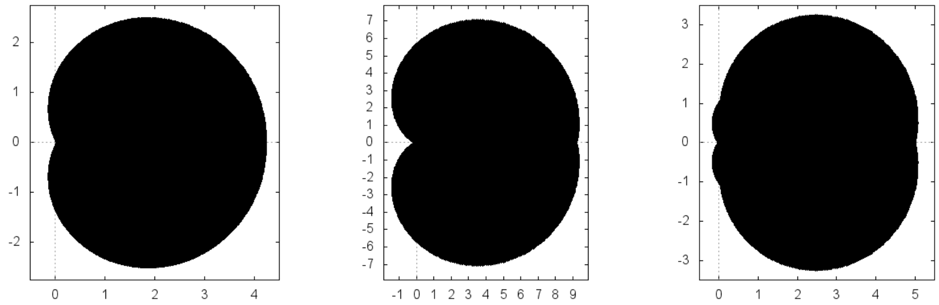
$$\rho_2 = \frac{e^{i\theta}-1}{3e^{i\theta}-1}$$

the absolute value of which attains its maximum for $\theta \in \{-\pi, \pi\}$ with the value ${}^1/_2$ which yields the bound $\beta = \arcsin({}^1/_2)$. To apply Lemma 3.9 we now need to establish γ for which we need to distinguish between the three considered values of δ:

- $\delta = 0$: We have $F(z) = {}^1/_2(1+z)$ which gives the bound $\gamma = \arcsin(1) = {}^\pi/_2$ so that Lemma 3.9 does not provide us with a positive angle α. However, Lemma 3.9 is only a sufficient condition and not necessary. Nevertheless, we can show that there is no positive angle such that the scheme is stable for all $\mu \in \mathcal{S}$ and $\lambda \in \mathcal{W}_\alpha$:

  From (3.14) we have

$$\varphi_\mu(z) = \frac{2(1-z)-\mu z(3-z)}{1+z}.$$



(a) SBDF2 (b) CNAB(1/16) (c) CNAB(1/4)

Figure 3.4.: Exterior of shaded region: Approximate stability domain for λ with arbitrary $\mu \in \mathcal{S}$ for the SBDF2, CNAB(1/16) and CNAB(1/4) respectively.

If we then take $z = -1 + i\varepsilon$ on the unit circle and $\mu = -1 - \varepsilon + \mathcal{O}(\varepsilon^2)$ on $\partial\mathcal{S}$ we get

$$\begin{aligned}\varphi_\mu(z) &= \frac{2(2-i\varepsilon) - (-1-i\varepsilon+\mathcal{O}(\varepsilon^2))(-1+i\varepsilon)(4-i\varepsilon)}{1+(-1+i\varepsilon)} \\ &= \frac{4-2i\varepsilon+(1+i\varepsilon+\mathcal{O}(\varepsilon^2))(4+5i\varepsilon+\varepsilon^2))}{i\varepsilon} \\ &= \frac{-i\varepsilon+\mathcal{O}(\varepsilon^2)}{i\varepsilon} \\ &= -1+\mathcal{O}(\varepsilon)\end{aligned}$$

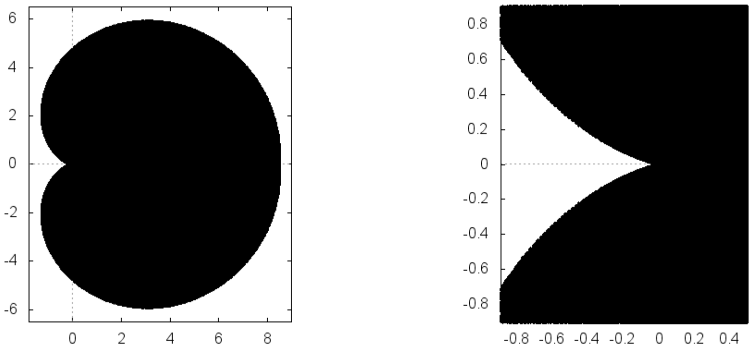
which shows that we can have values for $\varphi_\mu(z)$ arbitrarily close to the negative real axis.

- $\delta = 1/16$: Here $F(z) = 1/16(z+3)^2$ which gives us the bound $\gamma = 2\arcsin(1/3)$ which in turn gives the angle
$$\alpha = \frac{\pi}{2} - \arcsin\frac{1}{2} - 2\arcsin\frac{1}{3} \approx 0.12\pi.$$
- $\delta = 1/4$: This leads to $F(z) = 1/4(z^2+3)$ and thus gives the bound $\gamma = \arcsin(1/3)$. We then get the stability angle
$$\alpha = \frac{\pi}{2} - \arcsin\frac{1}{2} - \arcsin\frac{1}{3} \approx 0.23\pi.$$

By plotting the set $\{\varphi_\mu(z) \ : \ \mu \in \mathcal{S},\ |z| < 1\}$ which is the complement of the region of the values for λ for which the scheme is stable with arbitrary $\mu \in \mathcal{S}$, [FHV97] established experimental values for the value α. These sets can be seen in Figure 3.4. The values for α as measured by [FHV97] can be seen in Table 3.2.

Scheme	SBDF2	CNAB(1/16)	CNAB(1/4)
α_{theo}	0.31π	0.12π	0.23π
α_{exp}	0.32π	0.14π	0.30π

Table 3.2.: Theoretical and experimental values for α, denoted by α_{theo} and α_{exp} respectively.



(a) The entire region

(b) Enlargement of the origin

Figure 3.5.: Exterior of shaded region: Approximate stability domain for λ with arbitrary $\mu \in \mathcal{S}$ for the SBDF3 scheme.

We observe for the CNAB(1/16) and SBDF2 schemes that the theoretical bounds are close to the experimental ones. However, for the CNAB(1/4) scheme we see that experimental bound is significantly larger than the theoretical one. Nevertheless, since the SBDF2 scheme permits the larges angle α, we expect this scheme to be the least restrictive with respect to stability.

3.3.2. 3rd Order Schemes

SBDF3. To establish the value for α for $A(\alpha)$-stability of the implicit scheme for arbitrary values $\mu \in \mathcal{S}$, we proceed as in [FHV97] by plotting the set $\{\varphi_\mu(z) \ : \ \mu \in \mathcal{S}, \ |z| < 1\}$ for sufficiently many $|z| < 1$ and $\mu \in \mathcal{S}$. In this case we sampled 2.7×10^7 point from this set. The result can be seen in Figure 3.5. Note that the non-smooth boundary of the domains is due to the finite nature of our sample. However, this does not effect our measurement of α, as we have sufficiently many points close to the origin for a good estimate of the angle α. By enlarging the area around the origin we can measure the experimental bound for the angle α. In this case we get an upper bound of $\alpha \approx 0.11\pi$. This is again significantly smaller that the angle of $\alpha = 0.48\pi$ [HW96] for the BDF3 method.

Remark 3.10 (Summary). The semi-discrete analysis of time stepping schemes using a scalar test problem as we have done in the preceding two sections only has a very limited application to the incompressible flow setting we require the methods for. The analysis stems from an ODE framework, while our application is a DAE system. Furthermore, as we have already stated, this analysis assumes the operator to be treated explicitly to be linear which is not the case in the Navier-Stokes setting.

Nevertheless, we can draw the following conclusions. For one, we have seen that the SBDF1 scheme has the largest stability region for the explicit part and no restriction to it in order to obtain A-stability in the implicit part. We can therefore expect this scheme to allow the largest time steps since it is with the choice of time step that we can force the eigenvalues to be inside the stability region. Within the regime of second order multistep schemes we can expect the SBDF2 method to allow the largest time steps since it has the largest stability region for the explicit part of the scheme and present the largest angle for $A(\alpha)$-stability. Furthermore, the analysis has shown that the CNLF scheme is only sensible in a setting with a skew symmetric convection operator which would result in purely imaginary eigenvalues.

4. IMEX Runge-Kutta Schemes

As in Chapter 3 this chapter will consider the time integration of a DAE system of the form

$$M\dot{\mathbf{u}}_h = A(\mathbf{u}_h, p_h) + C(\mathbf{u}_h) \tag{4.1}$$

where M is the mass-matrix, A the Stokes operator and right-hand side while C is the convective part. Similar to the IMEX multistep methods we want to split the time-integration of the problem (4.1) into two parts treating $A(\mathbf{u}, p)$ implicitly and $C(\mathbf{u})$ explicitly. However we now do this within a Runge-Kutta setting. IMEX Runge-Kutta methods were introduced by [ARS97] and the underlying idea is to combine an explicit Runge-Kutta method with an implicit Runge-Kutta method which work on the same intermediate time-levels into a scheme which is expected to have the same accuracy as the minimum accuracy of the two separate methods.

The use of Runge-Kutta methods also presents the possibility of going to higher order methods compared to the use of backwards-difference formulas which become unstable if the order is greater than six. Indeed, [KC03] present methods up to order five for ODE systems originating from convection-diffusion-reaction equations.

We will begin with a short introduction to the basic definitions connected to Runge-Kutta methods. We then define IMEX Runge-Kutta schemes available in the literature suitable for our application which go up to formal order three. Finally, we consider the joint stability region of the resulting method.

4.1. Basics of Runge-Kutta Schemes

In this section we will introduce some of the basic notation and definitions regarding Runge-Kutta schemes, found, e.g. in [HNW93; HW96].

Definition 4.1. Let $s \in \mathbb{N}$, $a_{ij} \in \mathbb{R}$ for $1 \leq j < i \leq s$, $b_i \in \mathbb{R}$ for $1 \leq j \leq s$ and $c_j = \sum_{j=1}^{i-1} a_{ij}$. The scheme

$$\begin{aligned} k_i &= f(t + \tau c_i, x_n + \tau \sum_{j=1}^{i-1} a_{ij} k_j) \qquad \text{for } i = 1, \ldots, s \\ x_{n+1} &= x_n + \tau \sum_{i=1}^{s} b_i k_i \end{aligned}$$

is called an *s-stage explicit Runge-Kutta method* (ERK) for the problem

$$\begin{aligned} \dot{x} &= f(t, x) \\ x(t_0) &= x_0. \end{aligned} \tag{4.2}$$

Remark 4.2. Runge-Kutta schemes can be summarised in a Butcher tableau:

$$\begin{array}{c|ccccc}
0 & & & & & \\
c_2 & a_{21} & & & & \\
c_3 & a_{31} & a_{32} & & & \\
\vdots & \vdots & \vdots & \ddots & & \\
c_s & a_{s1} & a_{s2} & \cdots & a_{s,s-1} & \\
\hline
 & b_1 & b_2 & \cdots & b_{s-1} & b_s
\end{array}$$

Definition 4.3. Let $s \in \mathbb{N}$, $a_{ij} \in \mathbb{R}$, $b_i \in \mathbb{R}$ for $1 \leq i, j \leq s$ and $c_j = \sum_{j=1}^{s} a_{ij}$. The scheme

$$\begin{aligned}
k_i &= x_n + f(t + \tau c_i, \tau \sum_{j=1}^{s} a_{ij} k_j) \qquad \text{for } i = 1, \dots, s \\
x_{n+1} &= x_n + \tau \sum_{i=1}^{s} b_i k_i
\end{aligned}$$

is called an *s-stage Runge-Kutta method* for the problem (4.2). If $a_{ij} = 0$ for $j > i$ we call it a *diagonally implicit Runge-Kutta method* (DIRK) and if additionally $a_{ii} = \gamma$ for $1 \leq i \leq s$ we call it a *singularly diagonal implicit Runge-Kutta method* (SDIRK).

These methods can again be summarised in a Butcher tableau, now with non-zero entries on and possibly above the diagonal of the matrix A.

Definition and Proposition 4.4. Let $A \in \mathbb{R}^{s \times s}$ and $b \in \mathbb{R}^s$ be the coefficients of an implicit Runge-Kutta method. The function

$$R(z) = \frac{\det(I - zA + z\mathbb{1}b^T)}{\det(I - zA)}$$

with $\mathbb{1} = (1, \dots, 1)^T$, is called the *stability function* of the corresponding RK method. It can be interpreted as the numerical solution of the test problem

$$\dot{y} = \lambda y, \quad y_0 = 1$$

with $z = \lambda \tau$ after one step of the method.

Definition 4.5. The set

$$\mathcal{S} = \{z \in \mathbb{C} \ : \ |R(z)| \leq 1\}$$

is called the *stability domain* of the method. As for the multistep case we call the method *A-stable* if $\mathcal{S} \subset \mathbb{C}^-$.

Definition 4.6. A Runge-Kutta method is called *L-stable* if it is A-stable and it additionally holds that

$$\lim_{z \to \infty} R(z) = 0$$

Proposition 4.7. *Let $A \in \mathbb{R}^{s \times s}$ and $b \in \mathbb{R}^s$ be an s-stage implicit A-stable Runge-Kutta method such that A is non-singular. If A, b satisfy one of the following conditions*

$$b_j = a_{sj} \qquad \textit{for } j = 1, \dots, s \tag{4.3}$$

$$b_1 = a_{i1} \qquad \textit{for } i = 1, \dots, s, \tag{4.4}$$

then the method is L-stable.

$$\begin{array}{c|ccccc|ccccc}
0 & 0 & & & & & 0 & & & & \\
c_1 & \widehat{a}_{21} & & & & & a_{21} & a_{22} & & & \\
c_2 & \widehat{a}_{31} & \widehat{a}_{32} & & & & a_{31} & a_{32} & a_{33} & & \\
\vdots & \vdots & \vdots & \ddots & & & \vdots & \vdots & \vdots & \ddots & \\
c_s & \widehat{a}_{s+1,1} & \widehat{a}_{s+1,2} & \cdots & \widehat{a}_{s+1,s} & 0 & a_{s+1,1} & a_{s+1,2} & a_{s+1,3} & \cdots & a_{s+1,s+1} \\
\hline
 & \shortparallel & \shortparallel & \cdots & \shortparallel & \shortparallel & \shortparallel & \shortparallel & \shortparallel & \cdots & \shortparallel \\
 & \widehat{b}_1 & \widehat{b}_2 & \cdots & \widehat{b}_s & \widehat{b}_{s+1} & b_1 & b_2 & b_3 & \cdots & b_{s+1}
\end{array}$$

Table 4.1.: IMEX Runge-Kutta Butcher tableau consisting of a compatible pair of Runge-Kutta schemes.

Remark 4.8. Methods which satisfy (4.3), are called *stiffly accurate*. Stiffly accurate Runge-Kutta methods are of special relevance in the context of the Navier-Stokes equations because (4.2) is a DAE system. Stiffly accurate methods are important when dealing with DEA's [HW96] since the application of non-stiffly accurate methods to DAE systems can lead to a reduction of the order of the method [KM06].

4.2. IMEX Runge-Kutta Schemes

Following [ARS97], let $s \in \mathbb{N}$ and consider a s-stage DIRK method with the coefficients $\widetilde{A} \in \mathbb{R}^{s \times s}$, $\widetilde{b} \in \mathbb{R}^s$ and $\widetilde{c} \in \mathbb{R}^s$. Further set $\sigma = s + 1$ and consider an $(s + 1)$-stage ERK scheme given by the coefficient set $\widehat{A} \in \mathbb{R}^{\sigma \times \sigma}$ and $\widehat{b} \in \mathbb{R}^\sigma$ such that $\widehat{c}^T = (0, \widetilde{c}^T)$. We can then extend the Butcher tableau of the implicit scheme into a tableau of a $(s+1)$-stage method $A \in \mathbb{R}^{\sigma \times \sigma}$, $b \in \mathbb{R}^\sigma$ and $c \in \mathbb{R}^\sigma$ where

$$A = \begin{pmatrix} 0 & 0 \\ 0 & \widetilde{A} \end{pmatrix}, \qquad b^T = (0, \widetilde{b}^T), \qquad c^T = (o, \widetilde{c}^T) = \widehat{c}^T$$

We also assume that both the explicit and implicit methods are *stiffly accurate* as a result of which we have that $\mathbf{u}^{n+1} = \tilde{\mathbf{u}}^{s+1}$. Since we solve for $\tilde{\mathbf{u}}^s$ in each stage, the divergence constrained will be enforced on each stage. Having that $\mathbf{u}^{n+1} = \tilde{\mathbf{u}}^{s+1}$ is therefore essential since we aim for $\mathbf{u}^{n+1}$ to be exactly divergence free rather than in an interpolated sense. Combining the two Butcher tableaus of the explicit and implicit methods we get a general IMEX-RK Butcher tableau as shown in Table 4.1. Note that in Table 4.1 we have not restricted $a_{i1} = 0$ for $2 \leq i \leq s + 1$ as in [ARS97] in order to include the type of IMEX-RK schemes introduced by [KC03] which use ESDIRK methods to discretise the implicit part, i.e. the first stage of the implicit method is an explicit one. The resulting methods then work as described in Algorithm 4.1. Note that if we are in the ARS regime and only consider DIRK methods for the implicit part rather than allowing ESDIRK methods, the evaluation in line 2 of Algorithm 4.1 is unnecessary.

4.2.1. Examples of Stiffly Accurate IMEX Runge-Kutta Schemes

We will now list the schemes available in the cited literature (although we do not claim this list to be extensive) which fit into our setting. Most IMEX-RK methods available in the literature only consider a stiffly accurate scheme in the implicit part and then construct the

```
Data: $\mathbf{u}^n$
Evaluate $C^1 = C(\mathbf{u}^n)$
Evaluate $S^1 = A(\mathbf{u}^n)$
for $i = 2$ to $s + 1$ do
    Solve $(M + \tau a_{ii} A)\tilde{\mathbf{u}}^i = M\mathbf{u}^n - \tau \sum_{j=1}^{i-1} \left\{ \widehat{a}_{ij} C^j + a_{ij} S^j \right\}$
    if $i < s + 1$ then
        Evaluate $C^i = C(\tilde{\mathbf{u}}^i)$
        Evaluate $S^i = A(\tilde{\mathbf{u}}^i)$
    end
end
Result: $\mathbf{u}^{n+1} = \tilde{\mathbf{u}}^{s+1}$
```

Algorithm 4.1.: IMEX Runge-Kutta time-integration using a stiffly accurate method.

$$\begin{array}{c|cc|cc}
0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 \\
\hline
 & 1 & 0 & 0 & 1
\end{array}$$

Table 4.2.: The Butcher tableau of the first order ARS(1,1,1) IMEX Runge-Kutta method.

explicit scheme such that $\widehat{b} = b$, see for example [ARS97; Bos09; BPR13; KC03; PR05] for such schemes. As these methods do not fit into our setting we will not list them here.

As in [PR00] where a number of these methods are listed, we will characterise the schemes by the initials of their authors and the triplet (s, σ, p) where s is the number of stages in the implicit method, σ is the number of stages in the explicit method and p is the order of the overall method.

First Order Methods. The only first order method which fits into our setting, is the *forward-backward Euler* method presented by [ARS97] which we will refer to as ARS(1,1,1). The Butcher tableau for this method can be seen in Table 4.2. Observe that the resulting method is identical to the IMEX multistep SBDF1 method and will therefore not be considered any further here.

Second Order Methods. Also from [ARS97] we have the ARS(2,2,2) method which takes a second order stiffly accurate SDIRK method and constructs a suitable explicit method such that $\widehat{b}_i = \widehat{a}_{s+1,i}$. The coefficients of this scheme can be seen in Table 4.3.

Another stiffly accurate method is the the LRR(3,2,2) scheme from [LRR00]. Although not originally named as an IMEX-RK method, it was put into the IMEX-RK framework by [PR00]. The corresponding Butcher tableau can be seen in Table 4.4. The implicit part of this method is not singularly diagonally implicit.

Furthermore, we have the second order BPR(4,4,2) method from [BPR17] which uses four stages in both the explicit and implicit part. This schemes is stiffly accurate and uses a SDIRK method in the implicit part. The coefficients of the scheme can be seen in Table 4.5.

$$\begin{array}{c|ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma & \gamma & 0 & 0 & 0 & \gamma & 0 \\ 1 & \delta & 1-\delta & 0 & 0 & 1-\gamma & \gamma \\ \hline & \delta & 1-\delta & 0 & 0 & 1-\gamma & \gamma \end{array} \qquad \gamma = 1 - \frac{\sqrt{2}}{2},\ \delta = 1 - \frac{1}{2\gamma}$$

Table 4.3.: Butcher tableau of the second order ARS(2,2,2) IMEX Runge-Kutta method.

$$\begin{array}{c|cccc|cccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 \\ 1/3 & 1/3 & 0 & 0 & 0 & 0 & 0 & 1/3 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 3/4 & 1/4 \\ \hline & 0 & 1 & 0 & 0 & 0 & 0 & 3/4 & 1/4 \end{array}$$

Table 4.4.: Butcher tableau of the second order LRR(3,2,2) IMEX-RK method.

Third Order Methods. The third order IMEX-RK scheme presented by [ARS97] requires four stages for both the implicit and explicit part. The coefficients of this scheme can be seen in Table 4.6. Here the diagonal coefficient was chosen to be rational as 1/2 and we again have a stiffly accurate SDIRK method in the implicit part and $\widehat{b_i} = \widehat{a}_{s+1,i}$.

Another third order method was devised by [BPR13]. Here we require five implicit stages but only three explicit stages. However, since the implicit scheme is an ESDIRK method, we only have to solve four linear systems per time step so the computational effort required for this method comparable to the ARS(4,4,3) scheme. The coefficient set of this method can be seen in Table 4.7. Note that even though the implicit scheme is an ESDIRK method, none of the stages are computed with an explicit method. As we can see in Algorithm 4.1, an ESDIRK method only includes an explicit evaluation of the Stokes part with the solution of the previous time step in the right-hand side. As a result the solution resulting from this method will only be pointwise divergence free if the initial condition is also pointwise divergence free.

Higher-Order Schemes. IMEX Runge-Kutta methods of order four and five consisting of six and eight stages respectively were derived in [KC03]. Like the BPR(5,3,3) method the implicit part is treated with ESDIRK schemes so, for example, only five linear systems have to be solved for the fourth order method. Although the implicit part of the schemes considered are L-stable, these schemes are constructed such that $b = \widehat{b}$ so the methods are unsuitable for our applications.

$$\begin{array}{c|ccccc|ccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 1/4 & 0 & 0 & 0 & 0 & 0 & 1/4 & 0 & 0 & 0 \\ 1/4 & 13/4 & -3 & 0 & 0 & 0 & 0 & 0 & 1/4 & 0 & 0 \\ 3/4 & 1/4 & 0 & 1/2 & 0 & 0 & 0 & 1/24 & 11/24 & 1/4 & 0 \\ 1 & 0 & 1/3 & 1/6 & 1/2 & 0 & 0 & 11/24 & 1/6 & 1/8 & 1/4 \\ \hline 1 & 0 & 1/3 & 1/6 & 1/2 & 0 & 0 & 11/24 & 1/6 & 1/8 & 1/4 \end{array}$$

Table 4.5.: Butcher tableau of the second order BPR(4,4,2) IMEX-RK method.

0	0	0	0	0	0	0	0	0	0	0
1/2	1/2	0	0	0	0	0	1/2	0	0	0
2/3	11/18	1/18	0	0	0	0	1/6	1/2	0	0
1/2	5/6	−5/6	1/2	0	0	0	−1/2	1/2	1/2	0
1	1/4	7/4	3/4	−7/4	0	0	3/2	−3/2	1/2	1/2
	1/4	7/4	3/4	−7/4	0	0	3/2	−3/2	1/2	1/2

Table 4.6.: The Butcher tableau of the third order ARS(4,4,3) method.

0	0	0	0	0	0	0	0	0	0	0
1	1	0	0	0	0	1/2	1/2	0	0	0
2/3	4/9	2/9	0	0	0	5/18	−1/9	1/2	0	0
1	1/4	0	3/4	0	0	1/2	0	0	1/2	0
1	1/4	0	3/4	0	0	1/4	0	3/4	−1/2	1/2
	1/4	0	3/4	0	0	1/4	0	3/4	−1/2	1/2

Table 4.7.: The Butcher tableau of the third order BPR(5,3,3) method.

Computational Efficiency. Unlike the ARS(2,2,2) scheme the LRR(3,2,2) scheme consists of a DIRK method for the implicit part rather than an SDIRK method. As a result the linear system which has to be solved in line 4 of Algorithm 4.1 is different for each stage. This reduces the efficiency of the method since we have to compute and store three different matrix inverses which in turn makes the scheme very memory intensive. We will therefore not consider this scheme further.

Furthermore, the BPR(4,4,2) method although using a SDIRK scheme in the implicit part, requires four stages in each time step, twice as many as ARS(2,2,2). It is therefore twice as computationally expensive compared to the ARS(2,2,2) method and we will also not consider this scheme any further.

Stability. Rather than considering the stability regions of the implicit and explicit part separately, we will consider the stability region of the combined scheme. The stability function of an IMEX Runge-Kutta scheme is given by

$$R(x+iy) = \frac{\det(\mathrm{Id} - xA - iy\widehat{A} + z\mathbb{1}b^T + iy\mathbb{1}\widehat{b}^T)}{\det(\mathrm{Id} - xA - iy\widehat{A})},$$

c.f. [CdFN01; KC03]. Here $x = \tau\alpha$ and $y = \tau\beta$, where $\alpha \in \mathbb{R}$ represents the eigenvalue of the implicit part and $\beta \in \mathbb{R}$ represents the eigenvalue of the explicit part resulting from a scalar test equation

$$\dot{x} = \alpha x + i\beta x$$

as in [ARW95; ARS97]. Note that this is a different and more restrictive setting compared with Chapter 3 where α, β were allowed to be complex valued.

The stability function of the ARS(2,2,2) method is then given by

$$R(x+iy) = \frac{1 + (1-2\gamma)x + \gamma(\delta-1)y^2 + i[y + \gamma(1-\delta-\gamma)xy]}{(1-\gamma x)^2}.$$

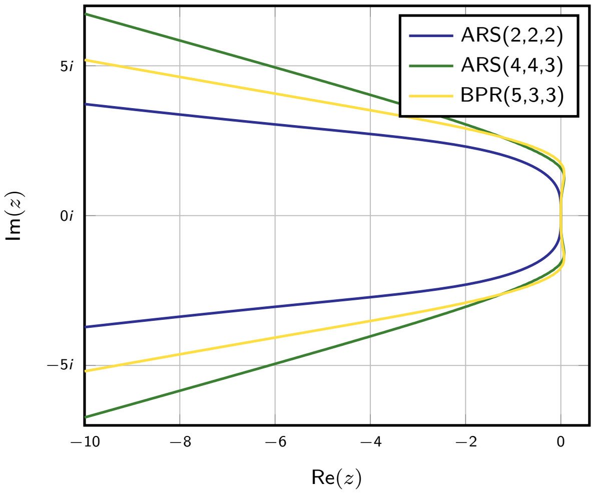


Figure 4.1.: Stability regions of the IMEX Runge-Kutta methods using an (E)SDIRK method in the implicit part.

The stability of the third order ARS(4,4,3) scheme is

$$R(x{+}iy) = \frac{48(6-6x+x^3)-144(1-x)y^2+3x^2y^2-7y^4+i[48(6-6x-y^2)y+29x^3y+57xy^3]}{18(x-2)^4}$$

and the stability function of the third order BPR(5,3,3) method is defined as

$$R(x{+}iy) = \frac{48-48x+8x^2-x^4-24y^2+24xy^2-4x^2y^2+i[48y-48xy+8x^3y-8y^3+8xy^3]}{3(x-2)^4}.$$

The boundary of the resulting stability domains, i.e. $|R(x+iy)| \leq 1$, can be seen in Figure 4.1. From this result we expect the ARS(4,4,3) to allow the largest time steps in practice while we expect the second order method to be the most restrictive. This compares with [ARS97] where the time step restriction on $y = \tau\beta$ was considered for a given ration $\alpha/\beta = x/y$.

5. Numerical Experiments

In this chapter we will use the IMEX schemes discussed in Chapters 3 and 4 for a number of computations in the open source high-order FEM package Netgen/NGSolve [Sch97; Sch14] available at `ngsolve.org`. We begin by discussing the inherent time step restriction arising from the explicit treatment of the convection operator. We then consider a number of numerical experiments both in 2 and 3 dimensions to illustrate the theoretical results and to evaluate the performance of the different schemes in practice.

5.1. Time Step Restrictions

5.1.1. IMEX Multistep Schemes

Explicit treatment of convection gives a restriction on the time step [KKW17; LS16]. For simplicity consider the scalar transport problem:

$$\begin{aligned} \partial_t u + a\partial_x u &= 0 \quad \text{in } (0,T]\times\Omega \\ u(0,\cdot) &= u_0 \quad \text{in } \Omega \end{aligned} \tag{5.1}$$

with some appropriate boundary condition. Finite element spatial semi-discretisation then leads to a system

$$\partial_t M u_h + \mathcal{L} u_h = 0. \tag{5.2}$$

Using an explicit scheme with constant time step τ to solve this in time then leads a scheme of the form

$$u_h^{n+1} = L u_h^n \tag{5.3}$$

where L is the resulting operator which advances the solution forward in time.

Definition 5.1 (*Stability*). For a given norm $\|\cdot\|$, a scheme of the form (5.3) is called *stable*, c.f. [HV03], if there exists a constant $C > 0$, independent of τ, such that

$$\|u_h^{n+1}\| \leq C\|u_h^0\|.$$

Lemma 5.2 (*Inverse estimate*). *Let $a < b \in \mathbb{R}$, $I = (a,b)$ and define $h = b - a$. For every polynomial $v \in \mathbb{P}^k(I)$ it holds that*

$$\|v'\|_{\mathcal{L}^2(I)} \leq 2\sqrt{3}\frac{k^2}{h}\|v\|_{\mathcal{L}^2(I)}\,. \tag{5.4}$$

Proof. See [Sch99, Theorem 3.91]. □

Lemma 5.3 (*Inverse trace estimate*). *In the context of Theorem 5.2 it holds*

$$\|v\|_{\mathcal{L}^2(\partial I)} \leq \frac{k+1}{\sqrt{h}}\|v\|_{\mathcal{L}^2(I)}. \tag{5.5}$$

Proof. See [WH03, Theorem 2]. □

Lemma 5.4 (*Time step restriction*). *An explicit time-stepping scheme applied to the semi-discretised transport problem* (5.2) *is stable provided the time step admits a* Courant-Friedrichs-Lewy *(CFL) condition of the form*

$$\tau \leq C \frac{h}{k^2} \frac{1}{|a|}$$

for some $C > 0$, independent of h, k and a.

Remark 5.5. The scaling of the CFL condition with respect to the polynomial order k is quadratic according to Lemma 5.4. For a pure transport problem this is a sharp bound [KS05] However, as we will see later, this bound does not seem to be sharp for the Navier-Stokes equations and numerically it has been observed in [FWK18; KKW17] that the scaling $k^{-1.5}$ seems to give a sharp bound for the CFL condition in the Navier-Stokes setting.

Proof of Lemma 5.4. We follow [HW08]. Let us consider an equidistant mesh $\{I^{\mathfrak{k}}\}$ of the interval (a, b) with $|I^{\mathfrak{k}}| = x_{\mathfrak{k}} - x_{\mathfrak{k}-1} = h$. Let $\{\ell_i^{\mathfrak{k}}\}_{i=1}^N$ be the local FEM basis on the element $I^{\mathfrak{k}}$. So on each element we have to solve

$$\partial_t M^{\mathfrak{k}} u_h + a C^{\mathfrak{k}} u_h = 0$$

with the local mass matrix

$$(M^{\mathfrak{k}})_{ij} = \int_{I^{\mathfrak{k}}} \ell_i^{\mathfrak{k}}(x) \ell_j^{\mathfrak{k}}(x) \,\mathrm{d}x$$

and local stiffness matrix

$$(C^{\mathfrak{k}})_{ij} = \int_{I^{\mathfrak{k}}} \ell_i^{\mathfrak{k}}(x) \frac{\mathrm{d}\ell_j^{\mathfrak{k}}}{\mathrm{d}x}(x) \,\mathrm{d}x.$$

Note that we omit any facet terms in the operator $C^{\mathfrak{k}}$. The addition of these would not effect the bound as volume derivatives will dominate due to the quadratic scaling in the inverse inequality (5.4) compared with the trace inequality (5.5). In [HW08] the operator $C^{\mathfrak{k}} - \mathcal{E}$ is considered where $\mathcal{E}$ is a zero matrix with unity entries in the diagonal corners, but this still gives a crude estimate and does not account for any choice of flux.

Now let $I = [-1, 1]$ be the reference domain and consider the affine transformation

$$x(r) = x_{\mathfrak{k}-1} + \frac{1+r}{2} h$$

from the reference domain to an element $I^{\mathfrak{k}}$. Furthermore let $\{\ell_i\}_{i=1}^N$ be the FEM basis on this reference domain. For the local mass matrix we then have

$$M_{ij}^{\mathfrak{k}} = \int_{I^{\mathfrak{k}}} \ell_i^{\mathfrak{k}}(x) \ell_j^{\mathfrak{k}}(x) \,\mathrm{d}x = \int_I \ell_i(r) \ell_j(r) \frac{h}{2} \,\mathrm{d}r = \frac{h}{2} M_{ij}$$

while for the stiffness matrix we observe

$$C_{ij}^{\mathfrak{k}} = \int_{I^{\mathfrak{k}}} \ell_i^{\mathfrak{k}}(x) \frac{\mathrm{d}\ell_j^{\mathfrak{k}}}{\mathrm{d}x}(x) \,\mathrm{d}x = \int_I \ell_i(r) \frac{\mathrm{d}\ell_j}{\mathrm{d}r}(r) \frac{2}{h} \frac{h}{2} \,\mathrm{d}r = C_{ij}$$

where the h scaling from the Jacobian of the integral transform is canceled out by the change of variable in the derivative term. On the reference domain we further introduce the differentiation matrix $\mathcal{D}_r$ defined by

$$(\mathcal{D}_r)_{ij} = \left.\frac{\mathrm{d}\ell_j}{\mathrm{d}r}\right|_{r_i}$$

where $\{r_m\}_{m=1}^N$ are the quadrature points on the reference interval, i.e, for a local approximation

$$u(r) \simeq u_h(r) = \sum_{m=1}^{N} u(r_m)\ell_m(r)$$

we have that

$$u_h'(r) = \mathcal{D}_r u_h(r).$$

For the mass, stiffness and differentiation matrices on the reference domain we observe

$$\begin{aligned}
(M\mathcal{D}_r)_{ij} &= \sum_{m=1}^{N} M_{im}\mathcal{D}_{r,mj} \\
&= \sum_{m=1}^{N} \int_I \ell_i(r)\ell_m(r)\,\mathrm{d}r \frac{\mathrm{d}\ell_j}{\mathrm{d}r}\bigg|_{r_m} \\
&= \int_I \ell_i(r) \sum_{m=1}^{N} \frac{\mathrm{d}\ell_j}{\mathrm{d}r}(r_m)\ell_m(r)\,\mathrm{d}r \\
&= \int_I \ell_i(r)\frac{\mathrm{d}\ell_j}{\mathrm{d}r}(r)\,\mathrm{d}r \\
&= C_{ij},
\end{aligned}$$

i.e.

$$M\mathcal{D}_r = C. \tag{5.6}$$

Having constructed the spatial operators let us consider the explicit Euler scheme for simplicity. For this we have

$$\begin{aligned}
&& M^{\mathfrak{k}}u_h^{n+1} - M^{\mathfrak{k}}u_h^n &= -\tau a C^{\mathfrak{k}}u_h^n \\
\iff && \frac{h}{2}Mu_h^{n+1} - \frac{h}{2}Mu_h^n &= -\tau a C u_h^n \\
\iff && Mu_h^{n+1} &= Mu_h^n - \frac{2\tau}{h}aCu_h^n \\
\iff && u_h^{n+1} &= u_h^n - \frac{2\tau}{h}aM^{-1}Cu_h^n \\
&& &= (\mathrm{Id} - \frac{2\tau}{h}aM^{-1}C)u_h^n.
\end{aligned}$$

So we have stability if $\|\,\mathrm{Id} - {}^{2\tau}/_{h}M^{-1}C\| \leq 1$. A necessary condition for this is

$$\frac{1}{h}\|aM^{-1}C\| \leq \frac{1}{\tau}.$$

Using (5.6) and the inverse estimate (5.4) we then have

$$\begin{aligned}
\|aM^{-1}C\|^2 &= \sup_{\|u_h\|_{\mathcal{L}^2}=1} \|aM^{-1}Cu_h\|_{\mathcal{L}^2}^2 \\
&= |a|^2 \sup_{\|u_h\|_{\mathcal{L}^2}=1} \|\mathcal{D}_r u_h\|_{\mathcal{L}^2}^2 \\
&= |a|^2 \sup_{\|u_h\|_{\mathcal{L}^2}=1} \|u_h'\|_{\mathcal{L}^2}^2 \\
&\leq |a|^2 \sup_{\|u_h\|_{\mathcal{L}^2}=1} C_1 k^4 \|u_h\|_{\mathcal{L}^2}^2.
\end{aligned}$$

Thus we have a necessary stability condition on the time step

$$\tau \leq C \frac{h}{k^2} \frac{1}{|a|}.$$

□

Remark 5.6. The same CFL scaling for a stable time step is obtained by [KS05], where an upper bound on the spectral radius of the time stepping operator is claimed, giving stability in the sense that the spectrum of the operator is inside the stability region of the explicit time stepping scheme c.f. [HV03, Chapter 6].

5.1.2. IMEX Runge-Kutta Schemes

The ARS(1,1,1) IMEX-RK method and the SBDF1 method are identical, which indicates that we should be faced with a similar time step restriction as for the IMEX multistep schemes.

Lemma 5.7 (*Time Step Restriction for Explicit Convection (Runge-Kutta)*). *Under the assumptions of Lemma 5.4 an explicit Runge-Kutta method for the model transport problem* (5.1) *is stable provided the CFL condition*

$$\tau \leq C_{con} \frac{h}{k^2} \frac{1}{|a|}$$

holds for some constant $C_{con} > 0$, independent of h and k.

Proof. Consider a two stage ERK method for the model transport problem (5.1). The method then reads

$$\begin{aligned}
c^1 &= C^{\mathfrak{k}}(\mathbf{u}^n) \\
\tilde{\mathbf{u}}^1 &= \mathbf{u}^n - \tau \widehat{a}_{21} (M^{\mathfrak{k}})^{-1} c^1 \\
c^2 &= C^{\mathfrak{k}}(\tilde{\mathbf{u}}^1) \\
\mathbf{u}^{n+1} &= \tilde{\mathbf{u}}^2 \\
&= \mathbf{u}^n - \tau(\widehat{a}_{31} (M^{\mathfrak{k}})^{-1} c^1 + \widehat{a}_{32} (M^{\mathfrak{k}})^{-1} c^2) \\
&= \mathbf{u}^n - \tau(\widehat{a}_{31} (M^{\mathfrak{k}})^{-1} C^{\mathfrak{k}}(\mathbf{u}^n) - \tau \widehat{a}_{32} (M^{\mathfrak{k}})^{-1} C(\mathbf{u}^n - \tau \widehat{a}_{21} (M^{\mathfrak{k}})^{-1} C^{\mathfrak{k}}(\mathbf{u}^n)) \\
&= \mathbf{u}^n - \tau(\widehat{a}_{31} + \widehat{a}_{32})(M^{\mathfrak{k}})^{-1} C^{\mathfrak{k}}(\mathbf{u}^n) + \tau^2 \widehat{a}_{31} \widehat{a}_{21} ((M^{\mathfrak{k}})^{-1} C^{\mathfrak{k}})^2 (\mathbf{u}^n) \\
&= \mathbf{u}^n - \tau (M^{\mathfrak{k}})^{-1} C^{\mathfrak{k}}(\mathbf{u}^n) + \frac{\tau^2}{2} ((M^{\mathfrak{k}})^{-1} C^{\mathfrak{k}})^2 (\mathbf{u}^n) \\
&= \mathbf{u}^n - \tau \frac{2}{h} M^{-1} C(\mathbf{u}^n) + \frac{\tau^2}{2} \left(\frac{2}{h} M^{-1} C \right)^2 (\mathbf{u}^n).
\end{aligned}$$

We have used the assumption $\widehat{a}_{si} = \widehat{b}_i$ on the one hand together with the first order condition $\sum_i \widehat{b}_i = 1$ to obtain

$$1 = \widehat{b}_1 + \widehat{b}_2 + \widehat{b}_3 = \widehat{a}_{31} + \widehat{a}_{32} + 0$$

and on the other with $\sum_j \widehat{a}_{ij} = c_i$ and the second order condition $\sum_i \widehat{b}_i \widehat{c}_i = 1/2$ to obtain

$$\frac{1}{2} = \widehat{b}_1 \cdot 0 + \widehat{b}_2 \cdot \widehat{c}_2 + 0 \cdot \widehat{c}_3 = \widehat{a}_{32} \cdot \widehat{a}_{21}.$$

In the final step we used that $M^{\mathfrak{k}} = \frac{h}{2}M$ and that $C^{\mathfrak{k}} = C$. Note, however, that the step of making $\mathbf{u}^{n+1}$ independent of the coefficients of the ERK scheme is not necessary since the coefficients $\widehat{a}_{ij}$ are real numbers independent of τ, h and k. So to obtain stability we essentially have the same bound (up to the constant) on the time step τ with respect to the norm of the operator $(M^{\mathfrak{k}})^{-1}C^{\mathfrak{k}} = \frac{2}{h}M^{-1}C$ giving again the CFL condition

$$\tau \leq C_{con}\frac{h}{k^2}.$$

□

5.2. Numerical Examples in Two Dimensions

5.2.1. General Set-up

In the following computations we will use $\boldsymbol{\mathcal{H}}(\mathrm{div})$-conforming $\mathbb{BDM}_k$ elements of order k for the velocity space $\mathbf{V}_h$ and piecewise polynomial, discontinuous elements $\mathbb{P}^{k-1}_{\mathrm{disc}}$ of order $k-1$ for the pressure space Q_h. As we have seen in Section 2.3, these two spaces build an inf-sup stable and pointwise divergence free finite element pair. This method will be abbreviated as BDMk.

In the SIP Stokes bilinear form we choose the jump penalisation parameter $\sigma = 6\frac{(k+1)(k+d)}{d}$ for the IMEX multistep schemes as used in [SL17a]. For the Runge-Kutta methods we use $\sigma = 4k^2$ as used by [SLL$^+$18]. Note that we have the essential scaling of k^2 for both choices, c.f. Remark 2.14. In the convective term we include the upwinding factor with $\gamma = 1$ as used in [SLL$^+$18].

To solve the arising linear systems $Ax = b$ we use a preconditioned Richardson iteration

$$x^{k+1} = x^k + P(b - Ax^k)$$

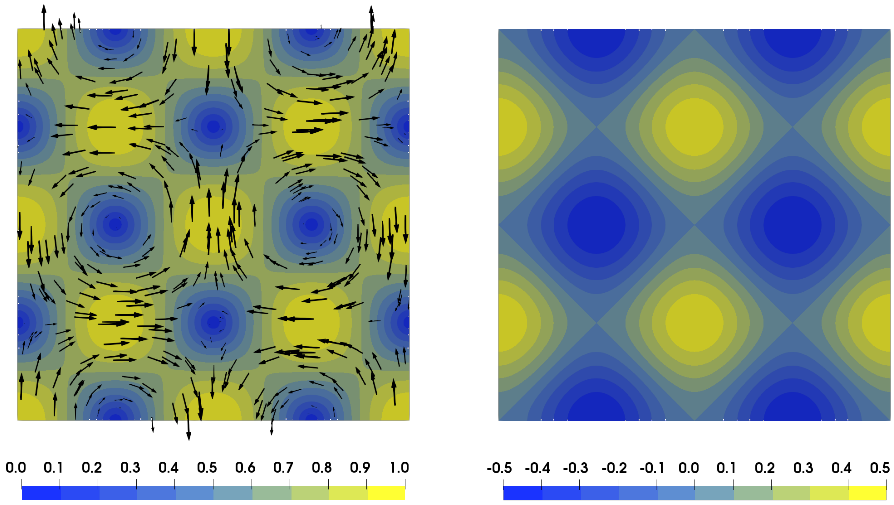
in each time step and we iterate until the relative ℓ^2-residual $\|b - Ax\|/\|b\|$ is less than 10^{-11}. The preconditioner P is the inverse of the pressure regularised system-matrix $M^*_{\text{p-reg}}$ where $M^*_{\text{p-reg}}(\mathbf{u}_h, p_h) = M^*(\mathbf{u}_h, p_h) - \rho\int_\Omega p_h q_h \,\mathrm{d}\mathbf{x}$ and the choice $\rho = 10^{-6}$. Since this matrix is symmetric and non-singular, we can compute the factorsation using `NGSolve`'s implementation of a sparse Cholesky factorisation which we found to be faster than using the direkt solver UMFPACK [Dav04].

5.2.2. Planar Lattice Flow

The first problem we will consider is the 'planar lattice flow' [Ber88] studied for example in [SL17a; SL17b; SLL$^+$18]. This initial velocity solves the stationary incompressible Euler equations [MB01] while in the viscous Navier-Stokes context diffusion and the time derivative balance. For $\nu \geq 0$ and $\mathbf{x} \in \Omega = (0,1)^2$ the exact solution is given by

$$\mathbf{u}(\mathbf{x}, t) = e^{-8\nu\pi^2 t}\begin{pmatrix}\sin(2\pi x_1)\sin(2\pi x_2)\\ \cos(2\pi x_1)\cos(2\pi x_2)\end{pmatrix}$$

$$p(\mathbf{x}, t) = \frac{1}{4}(\cos(4\pi x_1) - \cos(4\pi x_2))e^{-16\nu\pi^2 t}$$



(a) Velocity field and magnitude

(b) Pressure

Figure 5.1.: Initial state of the planar vortex problem.

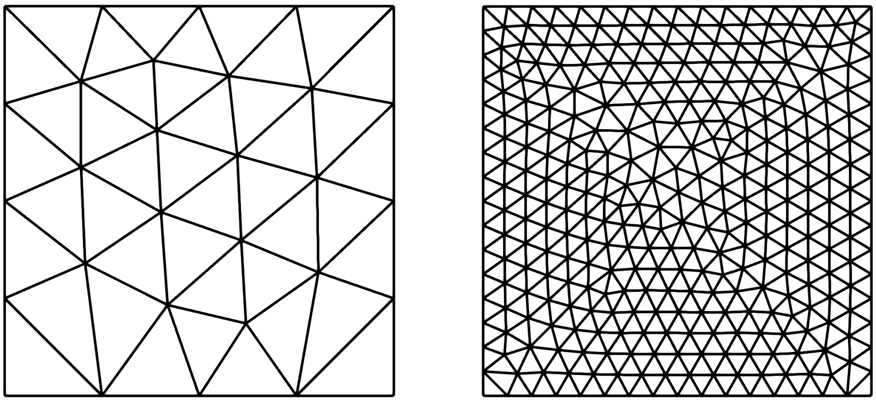


Figure 5.2.: Meshes with $h_{max} = 0.25$ resulting in 42 triangles and $h_{max} = 0.063$ resulting in 600 triangles respectively.

This flow has a saddle point structure and is therefore "dynamically unstable so that small perturbations result in a very chaotic motion" [MB01]. The velocity field and the pressure at $t = 0$ can be seen in Figure 5.1.

The two meshes used for the majority of the computations of this problem are shown in Figure 5.2. We will refer to the mesh with $h_{max} = 0.25$ as the *coarse* mesh and the mesh with $h_{max} = 0.063$ as the *fine* mesh. On the boundary we impose periodic boundary conditions. This ensures that the velocity is not artificially kept stable by fixing the FEM solution to the exact solution by a Dirichlet condition on (at least a part of) the boundary. The viscosity is chosen to be $\nu = 10^{-5}$. With the characteristic length $L = 1$ and the characteristic velocity $V = 1$ being the maximal velocity at $t = 0$, the Reynolds number is then $Re = 10^5$.

Time Step Restriction: Polynomial Order. As a result of Lemma 5.4, we expect that the explicit treatment of the convective term leads to a time restriction which scales linearly with

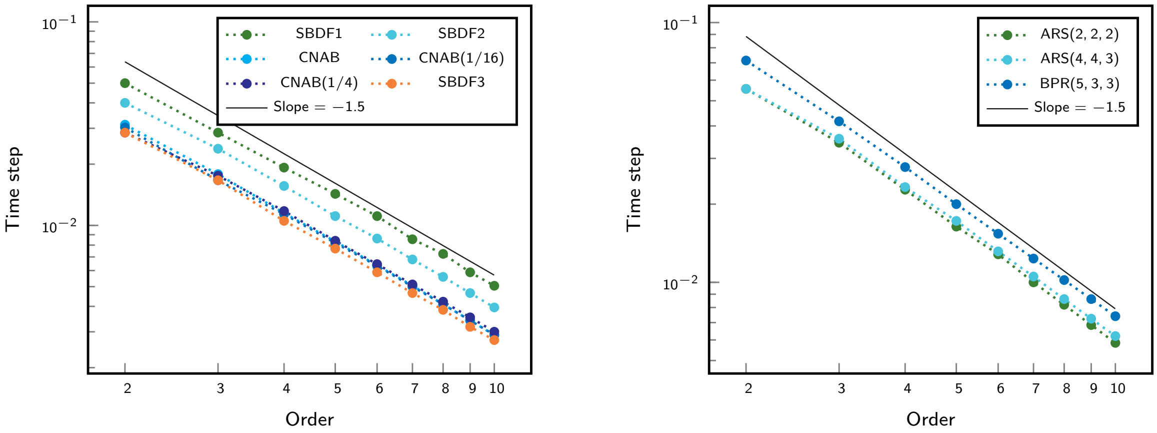


Figure 5.3.: Largest time step for a stable solution at $T = 1$ for different IMEX schemes and polynomial orders on the coarse mesh.

respect to the mesh resolution and magnitude of the convective velocity and scales quadratically with respect to the polynomial order, i.e.

$$\tau \leq C_{conv} \frac{h}{k^2} \frac{1}{\|\mathbf{u}\|_\infty}.$$

In order to establish the stability of the schemes considered in practice and to determine whether the above CFL condition is sharp we test to find the largest stable time step for a given mesh size h and polynomial order k. We consider the scheme to be stable if the numerical solution at $T = 1$ still has the same vortex structure as the initial solution.

In Figure 5.3 we see the results for all IMEX methods on the coarse mesh. We observe that the IMEX multistep method allowing the largest time step is the SBDF1 scheme followed by the SBDF2 scheme. The Crank-Nicolson schemes were the most restrictive within the second order schemes while the third order SBDF3 scheme had the strongest time step restriction of all considered schemes. The fact that the SBDF2 method is less restrictive compared with the CNAB(δ) schemes is consistent with the results from Chapter 3 where the SBDF2 scheme presented the largest stability area of the explicit part of the scheme and the largest A(α)-stability angle. Within the Runge-Kutta methods the BPR(5,3,3) scheme allowed the largest time steps while both ARS schemes had very similar stability limits. This is not consistent with the stability area computed in Section 4.2 from which we expected the ARS(4,4,3) method to allow the largest time step.

We note that rather than the quadratic scaling with respect to k an increase in polynomial order only required a time step decrease by a factor of about 3/2. This weaker restriction on the time step with respect to the polynomial order of the finite element space was also observed in [FWK18; KKW17] for the Navier-Stokes equation where an experimental factor of 1.5 was established for a method similar to SBDF2 from [KIO91] which treats the convective part explicitly but also splits the pressure from the diffusion term into a pressure Poisson equation. The computations in [FWK18] suggest that this is a sharp bound.

Time Step Restriction: Mesh Size. Figure 5.4 shows the largest "stable" time step for a series of finer meshes using the BDM8 elements and the set of IMEX schemes considered. The term "stable" is interpreted as before. We note that the second and third order schemes

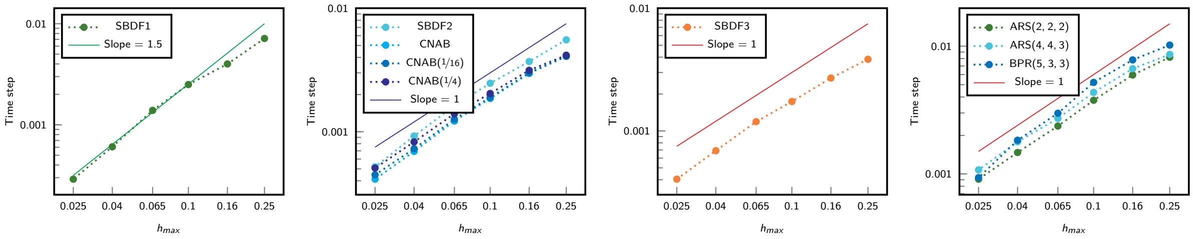


Figure 5.4.: Largest time step for a stable solution at $T = 1$ for different IMEX schemes using BDM8 elements.

performed as expected with a linear scaling between the mesh diameter and the largest stable time step. Within the multistep regime the SBDF2 scheme allowed for the largest time steps and the SBDF3 scheme performed similarly to the second order schemes. The IMEX Runge-Kutta methods performed similarly, with the BPR(5,3,3) method initially again allowing the largest time steps.

However, the first order SBDF1 scheme did not perform as expected. Here the scaling was closer to 3/2. Also on the finest mesh considered $h_{max} = 0.025$ the SBDF1 scheme has the strongest time step restrictions of all the schemes considered. It is unclear why this method scales differently with respect to h_{max}, and the analysis considered in Section 3.2 and Section 5.1 does not help since it only gave us heuristic information on the considered schemes.

For each time step of the ARS$(2, 2, 2)$ method we have to solve two linear systems and for the ARS$(4, 4, 3)$ and BPR$(5, 3, 3)$ methods we have to solve four linear systems in each time step respectively. This compares to the single linear system which has to be solved in each time step when using IMEX multistep methods. We therefore want to establish if and by how much IMEX Runge-Kutta methods allow for larger time steps to offset the larger number of linear systems which have to solved in a computation.

In Table 5.1 we see the reciprocal of the largest stable time step of the SBDF2 and SBDF3 IMEX multistep methods and of all three IMEX Runge-Kutta methods considered here for spatial orders $k = 2, \dots, 10$. Comparing the results for the IMEX Runge-Kutta methods and SBDF method of the same expected temporal order we observe that the second order ARS$(2, 2, 2)$ method admits a time step which is on average 1.46 times larger than that admitted by the SBDF2 method for a given polynomial order in the CFL limit. So the additional effort required to solve two linear systems instead of one is only partially offset by the possibility of a larger time step. We get a similar result when considering the third order schemes. The ARS$(4, 4, 3)$ method allows a time step which is on average 2.20 larger than for the SBDF3 method and the BPR$(5, 3, 3)$ method allows a time step which is on average 2.62 times larger than for the SBDF3 method for a give order k in the CFL stability limit. As a result the computational effort of having to solve four linear systems instead of one can again only be partially offset by the use of larger time steps.

In Table 5.2 we compare the largest stable time step of the same methods for different meshes. We observe that the ARS$(2, 2, 2)$ method admits a time step which is on average 1.57 times larger than that admitted by the SBDF2 method for a given mesh. So again, we cannot completely offset the additional effort of having to solve two linear systems instead of one by the choice of a larger time step in the CFL stability limit. For the third order Runge-Kutta methods we also see that the additional effort of having to solve four linear systems in each time step is only partially offset in the stability limit. The ARS$(4, 4, 3)$ method allows a time step which is on average 2.46 compared with SBDF3 and the BPR$(5, 3, 3)$ method allows a

	$1/\tau$					$\tau_{\text{RK}}/\tau_{\text{mult}}$		
Order	SBDF2	SBDF3	ARS(2, 2, 2)	ARS(4, 4, 3)	BPR(5, 3, 3)	ARS(2, 2, 2)/SBDF2	ARS(4, 4, 3)/SBDF3	BPR(5, 3, 3)/SBDF3
2	25	35	18	18	14	1.38889	1.94444	2.50000
3	42	60	29	28	24	1.44828	2.14286	2.50000
4	64	95	44	43	36	1.45455	2.20930	2.63889
5	90	130	61	58	50	1.47541	2.24138	2.60000
6	116	170	78	76	65	1.48718	2.23684	2.61538
7	147	215	100	95	81	1.47000	2.26316	2.65432
8	179	260	122	116	98	1.46721	2.24138	2.65306
9	215	315	146	138	116	1.47260	2.28261	2.71552
10	252	366	171	161	135	1.47368	2.27329	2.71111
Average						1.45976	2.20392	2.62092

Table 5.1.: Reciprocal of the largest stable time step for different polynomial orders on the coarse mesh and the factor by which IMEX Runge-Kutta methods allow a larger time step for a scheme of the same order compared with SBDF schemes of the same order.

	$1/\tau$					$\tau_{\text{RK}}/\tau_{\text{mult}}$		
h_{max}	SBDF2	SBDF3	ARS(2, 2, 2)	ARS(4, 4, 3)	BPR(5, 3, 3)	ARS(2, 2, 2)/SBDF2	ARS(4, 4, 3)/SBDF3	BPR(5, 3, 3)/SBDF3
0.25	180	260	122	116	98	1.47541	2.24138	2.65306
0.16	270	370	168	150	128	1.60714	2.46667	2.89063
0.1	405	575	265	230	192	1.5283	2.5	2.99479
0.065	615	840	422	368	335	1.45735	2.28261	2.50746
0.04	1084	1450	680	560	545	1.59412	2.58929	2.66055
0.025	1928	2475	1100	930	1070	1.75273	2.66129	2.31308
Average						1.56917	2.45687	2.66993

Table 5.2.: Reciprocal of the largest stable time step for different mesh diameters and BDM8 elements on the coarse mesh and the factor by which IMEX Runge-Kutta methods allow a larger time step for a scheme of the same order compared with SBDF schemes of the same order.

time step which is on average 2.69 larger than that for SBDF3 for a given mesh.

Note that this comparison in efficiency of these time stepping schemes is only valid in the CFL stability limit. However, we consider this to be a valid comparison since we do not observe any temporal discretisation error from the schemes with time steps below the CFL limit in this example, c.f. below.

Time Step Convergence. We aim to check that the IMEX methods considered are of the order with respect to the time step as expected from their construction. To ensure a high spatial resolution we use high order BDM8 elements on both the coarse and fine meshes. The results for the IMEX multistep methods and IMEX Runge-Kutta methods can be seen in Figure 5.5 and Figure 5.6 respectively. The dotted lines with circular mark indicate the results from the coarser mesh and the dashed lines with square mark indicate the results from the fine mesh. Note that it is sufficient to look at the error at $T = 1$, since the $\mathcal{L}^2$- and $\mathcal{H}^1$-errors are monotonically increasing in this example.

In each case the largest time step is very close to the stability limit, and the observed error for these time steps indicates that the solution is close to becoming unstable. On the fine mesh, we observe that the SBDF1 scheme is of order 1 as expected. We also see that the stability restriction on the time step for all second and third order schemes is so strong, that the dominating source of error is the spatial discretisation. This is indicated by the fact that after an initial sharp drop in the error there is no improvement for smaller time steps. The fact that second order time stepping is sufficient in this example is not surprising since in time

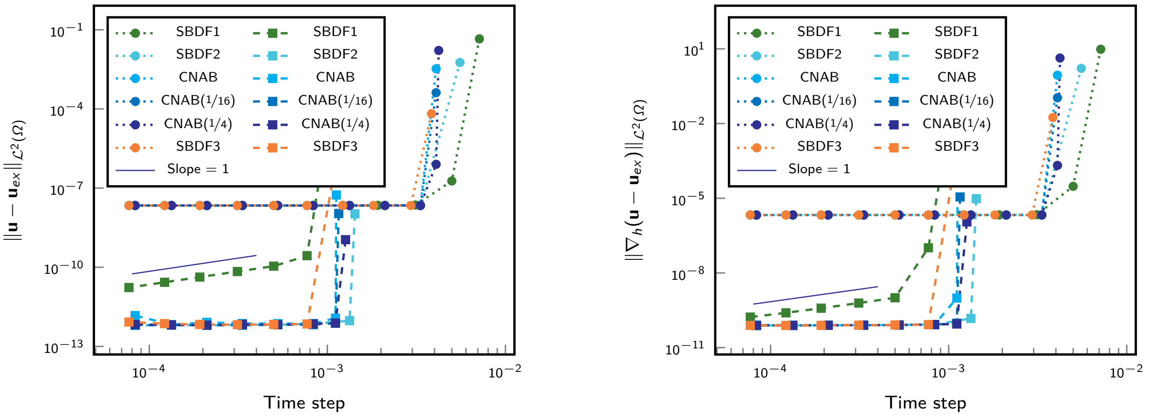


Figure 5.5.: The $\mathcal{L}^2$-norm and $\mathcal{H}^1$-semi-norm velocity errors at $T = 1$, computed using BDM8 elements and IMEX multistep methods. Computations on the coarse mesh ($h_{max} = 0.25$) are indicated by dotted lines with round markings while the computations using the fine mesh ($h_{max} = 0.063$) are indicated by dashed lines with square markings.

h_{max}	$\|\mathbf{u} - \mathbf{u}_{ex}\|_{\mathcal{L}^2(\Omega)}$							$\|\nabla_h(\mathbf{u} - \mathbf{u}_{ex})\|_{\mathcal{L}^2(\Omega)}$						
	SBDF1	SBDF2	CNAB	CNAB(1/16)	CNBA(1/4)	SBDF3	Order	SBDF1	SBDF2	CNAB	CNAB(1/16)	CNBA(1/4)	SBDF3	Order
0.25	3.19×10^{-2}	3.19×10^{-2}	3.19×10^{-2}	3.19×10^{-2}	3.19×10^{-2}	3.19×10^{-2}	-	1.10	1.10	1.10	1.10	1.10	1.10	-
0.16	1.16×10^{-2}	1.16×10^{-2}	1.16×10^{-2}	1.16×10^{-2}	1.16×10^{-2}	1.16×10^{-2}	2.3	4.84×10^{-1}	4.84×10^{-1}	4.84×10^{-1}	4.84×10^{-1}	4.84×10^{-1}	4.84×10^{-1}	1.8
0.1	2.82×10^{-3}	2.82×10^{-3}	2.82×10^{-3}	2.82×10^{-3}	2.82×10^{-3}	2.82×10^{-3}	3.0	2.04×10^{-1}	2.04×10^{-1}	2.04×10^{-1}	2.04×10^{-1}	2.04×10^{-1}	2.04×10^{-1}	1.8
0.063	4.35×10^{-4}	4.35×10^{-4}	4.35×10^{-4}	4.35×10^{-4}	4.35×10^{-4}	4.35×10^{-4}	4.0	7.46×10^{-2}	7.46×10^{-2}	7.46×10^{-2}	7.46×10^{-2}	7.46×10^{-2}	7.46×10^{-2}	2.2
0.04	9.21×10^{-5}	9.21×10^{-5}	9.21×10^{-5}	9.21×10^{-5}	9.21×10^{-5}	9.21×10^{-5}	3.4	2.83×10^{-2}	2.83×10^{-2}	2.83×10^{-2}	2.83×10^{-2}	2.83×10^{-2}	2.83×10^{-2}	2.1
0.025	2.32×10^{-5}	2.32×10^{-5}	2.32×10^{-5}	2.32×10^{-5}	2.32×10^{-5}	2.32×10^{-5}	2.9	1.15×10^{-2}	1.15×10^{-2}	1.15×10^{-2}	1.15×10^{-2}	1.15×10^{-2}	1.15×10^{-2}	1.9

Table 5.3.: The $\mathcal{L}^2$-norm and $\mathcal{H}^1$-semi-norm velocity error at $T = 1$, computed using BDM2 elements and the constant time step $\tau = 10^{-4}$. The convergence orders are computed from the SBDF2 results.

the solution only changes with a mild exponential factor due to the large Reynolds number chosen in this experiment.

Space Convergence. To see if the IMEX schemes have an effect on convergence with respect to mesh resolution we consider the $\mathcal{L}^2$-norm and $\mathcal{H}^1$-semi-norm error at $T = 1$ for a series of meshes with h_{max} between 0.25 and 0.025 using BDM2, BDM4 and BDM8 elements with a constant time step $\tau = 10^{-4}$.

In Figure 5.7 we see the results for the SBDF2 scheme and the convergence rates are better than expected from Corollary 2.25. We observe convergence of order $k+1$ in the $\mathcal{L}^\infty(0, T; \boldsymbol{\mathcal{L}}^2)$-norm and of order k in the $\mathcal{L}^\infty(0, T; \boldsymbol{\mathcal{H}}^1)$-norm. Figure 5.8 shows the results for the ARS(2,2,2) scheme. For order 8 elements we do not see any improvement in the error for meshes smaller that $h_{max} = 0.063$, since the error is at machine precision. Table 5.3, Table 5.4 and Table 5.5 show the results for all IMEX multistep methods considered while Table 5.6, Table 5.7 and Table 5.8 show the results from the IMEX Runge-Kutta methods. For order 2 and 4 all schemes give the same results. For order 8 we see that the SBDF1 scheme gives larger errors. As we have seen above this is since with $\tau = 10^{-4}$, $k = 8$ and $h_{max} = 0.063$, the time discretisation error dominates over the spatial error.

Large T. To see how the different schemes perform when considering larger and practically relevant end times T we take $T = 40$. For the spatial discretisation we use BDM8 elements on the fine mesh. For the temporal discretisation we choose two constant time steps $\tau = 10^{-3}$

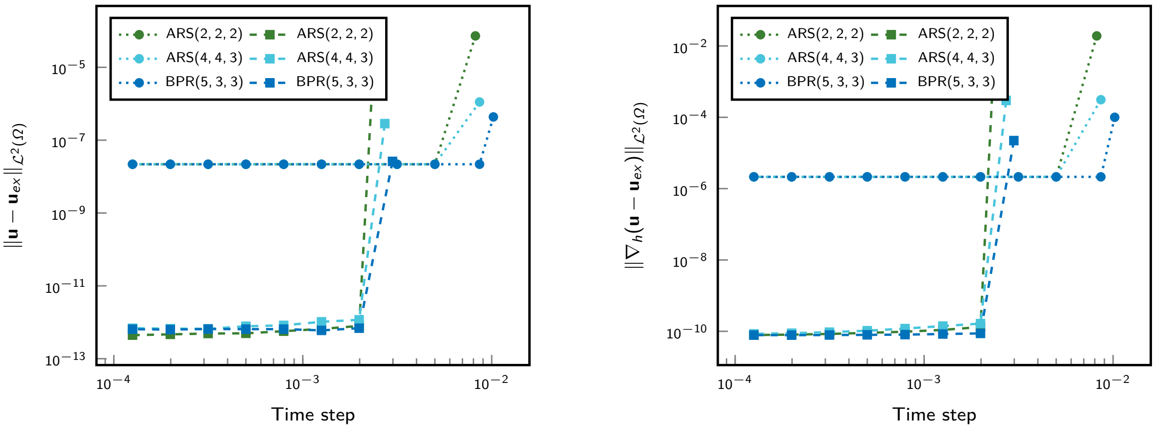


Figure 5.6.: The $\mathcal{L}^2$-norm and $\mathcal{H}^1$-semi-norm velocity errors at $T = 1$, computed using BDM8 elements and IMEX Runge-Kutta methods. Computations on the coarse mesh ($h_{max} = 0.25$) are indicated by dotted lines with round markings while the computations using the fine mesh ($h_{max} = 0.063$) are indicated by dashed lines with square markings.

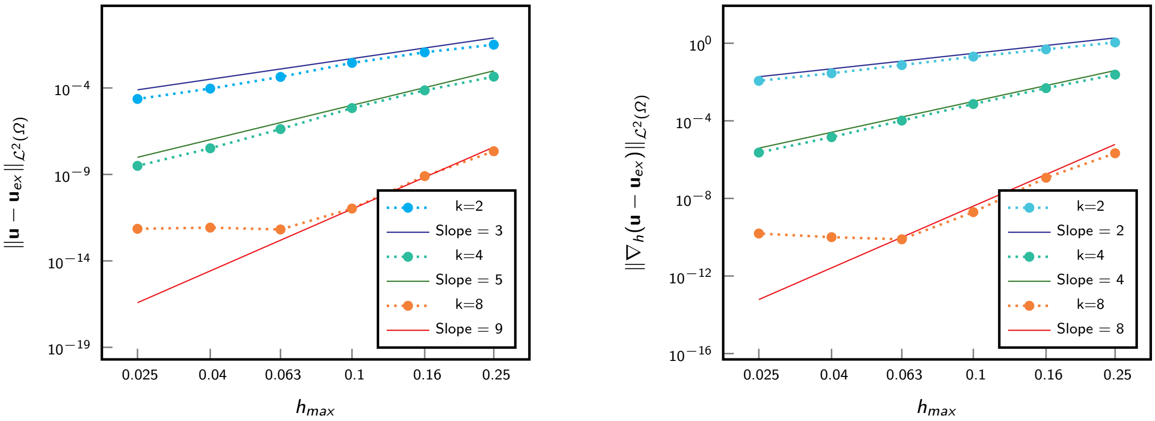


Figure 5.7.: Mesh convergence: The $\mathcal{L}^2$-norm and $\mathcal{H}^1$-semi-norm velocity error at $T = 1$, computed with BDMk elements on the coarse mesh using the SBDF2 time stepping method.

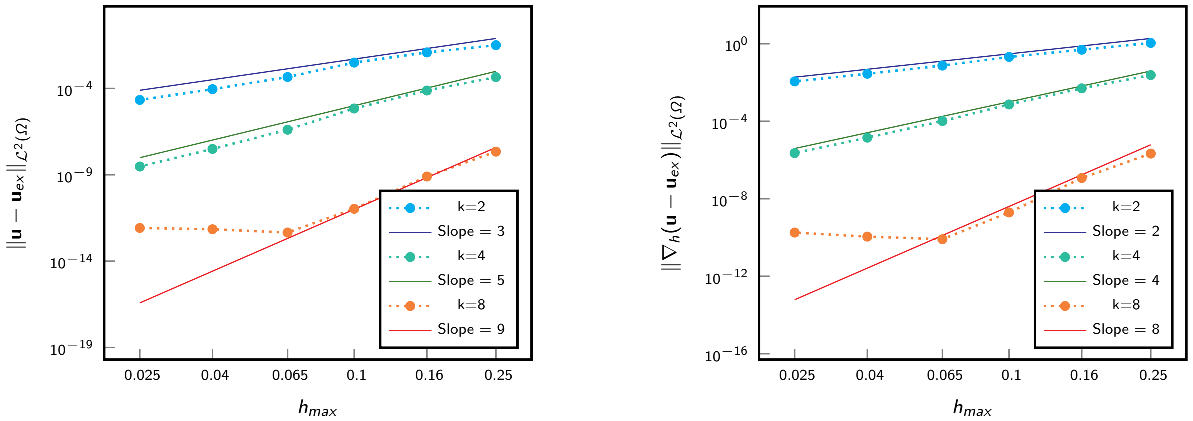


Figure 5.8.: Mesh convergence: The $\mathcal{L}^2$-norm and $\mathcal{H}^1$-semi-norm velocity error at $T = 1$, computed with BDMk elements on the coarse mesh and using the ARS$(2, 2, 2)$ IMEX Runge-Kutta time stepping method.

h_{max}	$\Vert\mathbf{u}-\mathbf{u}_{ex}\Vert_{\mathcal{L}^2(\Omega)}$						$\Vert\nabla_h(\mathbf{u}-\mathbf{u}_{ex})\Vert_{\mathcal{L}^2(\Omega)}$							
	SBDF1	SBDF2	CNAB	CNAB(1/16)	CNBA(1/4)	SBDF3	Order	SBDF1	SBDF2	CNAB	CNAB(1/16)	CNBA(1/4)	SBDF3	Order
0.25	4.50×10^{-4}	4.50×10^{-4}	4.50×10^{-4}	4.50×10^{-4}	4.50×10^{-4}	4.50×10^{-4}	-	2.41×10^{-2}	2.41×10^{-2}	2.41×10^{-2}	2.41×10^{-2}	2.41×10^{-2}	2.41×10^{-2}	-
0.16	7.26×10^{-5}	7.26×10^{-5}	7.26×10^{-5}	7.26×10^{-5}	7.26×10^{-5}	7.26×10^{-5}	4.1	4.82×10^{-3}	4.82×10^{-3}	4.82×10^{-3}	4.82×10^{-3}	4.82×10^{-3}	4.82×10^{-3}	3.6
0.1	6.86×10^{-6}	6.86×10^{-6}	6.86×10^{-6}	6.86×10^{-6}	6.86×10^{-6}	6.86×10^{-6}	5.0	7.32×10^{-4}	7.32×10^{-4}	7.32×10^{-4}	7.32×10^{-4}	7.32×10^{-4}	7.32×10^{-4}	4.0
0.063	4.19×10^{-7}	4.19×10^{-7}	4.19×10^{-7}	4.19×10^{-7}	4.19×10^{-7}	4.19×10^{-7}	6.1	1.03×10^{-4}	1.03×10^{-4}	1.03×10^{-4}	1.03×10^{-4}	1.03×10^{-4}	1.03×10^{-4}	4.2
0.04	3.25×10^{-8}	3.24×10^{-8}	3.24×10^{-8}	3.24×10^{-8}	3.24×10^{-8}	3.24×10^{-8}	5.6	1.44×10^{-5}	1.44×10^{-5}	1.44×10^{-5}	1.44×10^{-5}	1.44×10^{-5}	1.44×10^{-5}	4.3
0.025	3.14×10^{-9}	3.13×10^{-9}	3.13×10^{-9}	3.13×10^{-9}	3.13×10^{-9}	3.13×10^{-9}	5.0	2.31×10^{-6}	2.31×10^{-6}	2.31×10^{-6}	2.31×10^{-6}	2.31×10^{-6}	2.31×10^{-6}	3.9

Table 5.4.: The $\mathcal{L}^2$-norm and $\mathcal{H}^1$-semi-norm velocity error at $T = 1$, computed with BDM4 elements and the constant time step $\tau = 10^{-4}$. The convergence orders are computed from the SBDF2 results.

h_{max}	$\Vert\mathbf{u}-\mathbf{u}_{ex}\Vert_{\mathcal{L}^2(\Omega)}$						$\Vert\nabla_h(\mathbf{u}-\mathbf{u}_{ex})\Vert_{\mathcal{L}^2(\Omega)}$							
	SBDF1	SBDF2	CNAB	CNAB(1/16)	CNBA(1/4)	SBDF3	Order	SBDF1	SBDF2	CNAB	CNAB(1/16)	CNBA(1/4)	SBDF3	Order
0.25	2.19×10^{-8}	2.19×10^{-8}	2.19×10^{-8}	2.19×10^{-8}	2.19×10^{-8}	2.19×10^{-8}	-	2.13×10^{-6}	2.13×10^{-6}	2.13×10^{-6}	2.13×10^{-6}	2.13×10^{-6}	2.13×10^{-6}	-
0.16	7.96×10^{-10}	7.96×10^{-10}	7.96×10^{-10}	7.96×10^{-10}	7.96×10^{-10}	7.96×10^{-10}	7.4	1.16×10^{-7}	1.16×10^{-7}	1.16×10^{-7}	1.16×10^{-7}	1.16×10^{-7}	1.16×10^{-7}	6.5
0.1	2.44×10^{-11}	1.06×10^{-11}	1.06×10^{-11}	1.06×10^{-11}	1.06×10^{-11}	1.06×10^{-11}	9.2	1.97×10^{-9}	1.96×10^{-9}	1.96×10^{-9}	1.96×10^{-9}	1.96×10^{-9}	1.96×10^{-9}	8.7
0.063	2.21×10^{-11}	6.55×10^{-13}	1.13×10^{-12}	6.40×10^{-13}	6.38×10^{-13}	7.85×10^{-13}	6.0	2.10×10^{-10}	7.78×10^{-11}	7.93×10^{-11}	7.80×10^{-11}	7.84×10^{-11}	7.82×10^{-11}	7.0
0.04	2.21×10^{-11}	8.40×10^{-13}	1.30×10^{-12}	8.47×10^{-13}	8.45×10^{-13}	9.47×10^{-13}	-0.5	2.21×10^{-10}	1.00×10^{-10}	1.01×10^{-10}	1.00×10^{-10}	1.00×10^{-10}	1.01×10^{-10}	-0.5
0.025	2.21×10^{-11}	7.16×10^{-13}	1.25×10^{-12}	7.30×10^{-13}	7.34×10^{-13}	8.56×10^{-13}	0.3	2.71×10^{-10}	1.55×10^{-10}	1.55×10^{-10}	1.55×10^{-10}	1.54×10^{-10}	1.55×10^{-10}	-0.9

Table 5.5.: The $\mathcal{L}^2$-norm and $\mathcal{H}^1$-semi-norm velocity error at $T = 1$, computed using BDM8 elements and the constant time step $\tau = 10^{-4}$. The convergence orders are computed from the SBDF2 results.

h_{max}	$\Vert\mathbf{u}-\mathbf{u}_{ex}\Vert_{\mathcal{L}^2(\Omega)}$				$\Vert\nabla_h(\mathbf{u}-\mathbf{u}_{ex})\Vert_{\mathcal{L}^2(\Omega)}$			
	ARS(2,2,2)	ARS(4,4,3)	BPR(5,3,3)	Order	ARS(2,2,2)	ARS(4,4,3)	BPR(5,3,3)	Order
0.25	3.19×10^{-2}	3.19×10^{-2}	3.19×10^{-2}	-	1.10	1.10	1.10	-
0.16	1.22×10^{-2}	1.22×10^{-2}	1.22×10^{-2}	2.2	4.91×10^{-1}	4.91×10^{-1}	4.91×10^{-1}	1.8
0.1	3.13×10^{-3}	3.13×10^{-3}	3.13×10^{-3}	2.9	2.08×10^{-1}	2.08×10^{-1}	2.08×10^{-1}	1.8
0.065	4.58×10^{-4}	4.58×10^{-4}	4.58×10^{-4}	4.5	7.46×10^{-2}	7.46×10^{-2}	7.46×10^{-2}	2.4
0.04	9.05×10^{-5}	9.05×10^{-5}	9.05×10^{-5}	3.3	2.82×10^{-2}	2.82×10^{-2}	2.82×10^{-2}	2.0
0.025	2.16×10^{-5}	2.16×10^{-5}	2.16×10^{-5}	3.0	1.13×10^{-2}	1.13×10^{-2}	1.13×10^{-2}	1.9

Table 5.6.: The $\mathcal{L}^2$-norm and $\mathcal{H}^1$-semi-norm velocity error at $T = 1$, computed using BDM2 elements and the constant time step $\tau = 2 \times 10^{-3}$. The convergence orders are computed from the ARS(2,2,2) results.

h_{max}	$\Vert\mathbf{u}-\mathbf{u}_{ex}\Vert_{\mathcal{L}^2(\Omega)}$				$\Vert\nabla_h(\mathbf{u}-\mathbf{u}_{ex})\Vert_{\mathcal{L}^2(\Omega)}$			
	ARS(2,2,2)	ARS(4,4,3)	BPR(5,3,3)	Order	ARS(2,2,2)	ARS(4,4,3)	BPR(5,3,3)	Order
0.25	4.51×10^{-4}	4.51×10^{-4}	4.51×10^{-4}	-	2.42×10^{-2}	2.42×10^{-2}	2.42×10^{-2}	-
0.16	7.57×10^{-5}	7.57×10^{-5}	7.57×10^{-5}	4.0	4.99×10^{-3}	4.99×10^{-3}	4.99×10^{-3}	3.5
0.1	6.88×10^{-6}	6.88×10^{-6}	6.88×10^{-6}	5.1	7.35×10^{-4}	7.35×10^{-4}	7.35×10^{-4}	4.1
0.065	4.07×10^{-7}	4.07×10^{-7}	4.07×10^{-7}	6.6	1.02×10^{-4}	1.02×10^{-4}	1.02×10^{-4}	4.6
0.04	3.14×10^{-8}	3.14×10^{-8}	3.14×10^{-8}	5.3	1.42×10^{-5}	1.42×10^{-5}	1.42×10^{-5}	4.1
0.025	3.00×10^{-9}	3.00×10^{-9}	3.00×10^{-9}	5.0	2.26×10^{-6}	2.26×10^{-6}	2.26×10^{-6}	3.9

Table 5.7.: The $\mathcal{L}^2$-norm and $\mathcal{H}^1$-semi-norm velocity error at $T = 1$, computed using BDM4 elements and the constant time step $\tau = 2 \times 10^{-3}$. The convergence orders are computed from the ARS(2,2,2) results.

h_{max}	$\Vert\mathbf{u}-\mathbf{u}_{ex}\Vert_{\mathcal{L}^2(\Omega)}$				$\Vert\nabla_h(\mathbf{u}-\mathbf{u}_{ex})\Vert_{\mathcal{L}^2(\Omega)}$			
	ARS(2,2,2)	ARS(4,4,3)	BPR(5,3,3)	Order	ARS(2,2,2)	ARS(4,4,3)	BPR(5,3,3)	Order
0.25	2.19×10^{-8}	2.19×10^{-8}	2.19×10^{-8}	-	2.14×10^{-6}	2.14×10^{-6}	2.14×10^{-6}	-
0.16	7.94×10^{-10}	7.94×10^{-10}	7.94×10^{-10}	7.4	1.17×10^{-7}	1.17×10^{-7}	1.17×10^{-7}	6.5
0.1	1.07×10^{-11}	1.07×10^{-11}	1.07×10^{-11}	9.2	1.96×10^{-9}	1.96×10^{-9}	1.96×10^{-9}	8.7
0.065	4.53×10^{-13}	5.12×10^{-13}	4.30×10^{-13}	7.3	8.10×10^{-11}	8.63×10^{-11}	7.65×10^{-11}	7.4
0.04	7.00×10^{-13}	7.58×10^{-13}	6.62×10^{-13}	-0.9	1.11×10^{-10}	1.25×10^{-10}	9.74×10^{-11}	-0.6
0.025	8.34×10^{-13}	8.61×10^{-13}	7.96×10^{-13}	-0.4	1.79×10^{-10}	2.10×10^{-10}	1.51×10^{-10}	-1.0

Table 5.8.: The $\mathcal{L}^2$-norm and $\mathcal{H}^1$-semi-norm velocity error at $T = 1$, computed using BDM8 elements and the constant time step $\tau = 2 \times 10^{-3}$. The convergence orders are computed from the ARS(2,2,2) results.

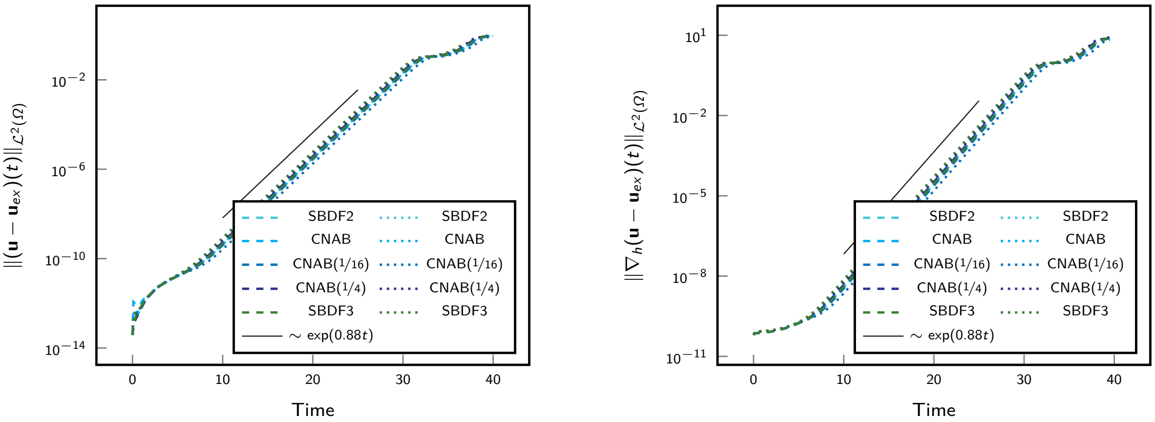


Figure 5.9.: Development of the $\mathcal{L}^2$-norm and $\mathcal{H}^1$-semi-norm velocity error on the fine mesh with BDM8 elements. The dotted lines indicate the time step $\tau = 10^{-3}$ and dashed lines $\tau = 10^{-4}$.

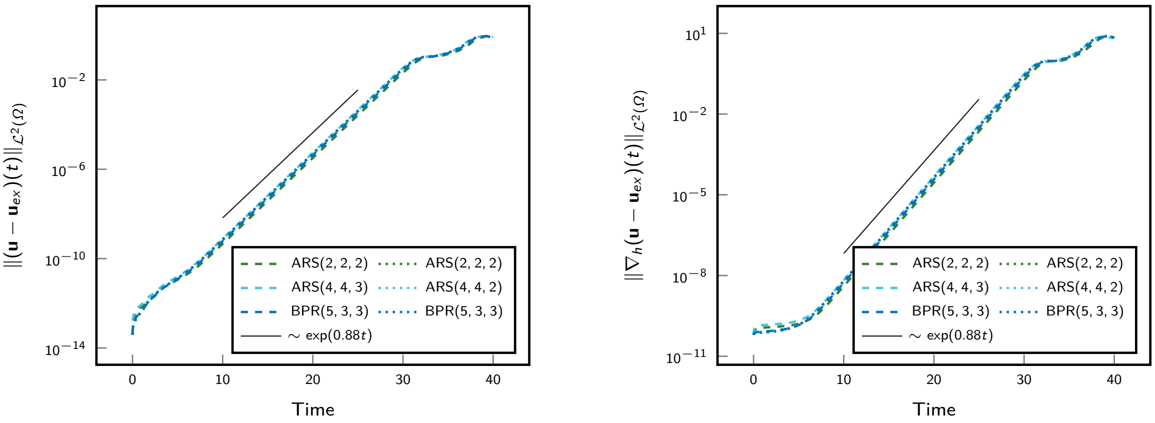


Figure 5.10.: The $\mathcal{L}^2$-norm and $\mathcal{H}^1$-semi-norm velocity error on the fine mesh with BDM8 elements. The dotted lines indicate the time step $\tau = 1/700$ and dashed lines $\tau = 1/7000$.

and $\tau = 10^{-4}$ for the multistep methods and for the Runge-Kutta methods we use time steps 1.4 times larger. This is because we established above that in the CFL stability limit the ARS(2,2,2) method allows time steps which are 1.4 times larger than the stable time steps for the SBDF2 method. Note that $\tau = 10^{-3}$ is close to the stability limit of the SBDF3 scheme on the fine mesh we established earlier.

The development of the $\mathcal{L}^2$-norm and $\mathcal{H}^1$-semi-norm velocity error over time for the IMEX multistep methods and IMEX Runge-Kutta methods can be seen in Figure 5.9 and Figure 5.10 respectively. We see that all IMEX schemes considered give us nearly identical results with both time steps. If we look at Figure 5.11, we can observe that the smallest error was obtained by the CNAB(1/16) scheme, however, the difference between the schemes is negligible. We also clearly observe the mild exponential growth in the error as expected from Theorem 2.24. The exponential growth stops at close to $T = 34$ where the FEM solution no longer has standing vortex structure of the exact solution and we see in Figure 5.12 that at $T = 40$ the four vertices have merged into two vertices and the solution no longer has the structure of the analytical solution.

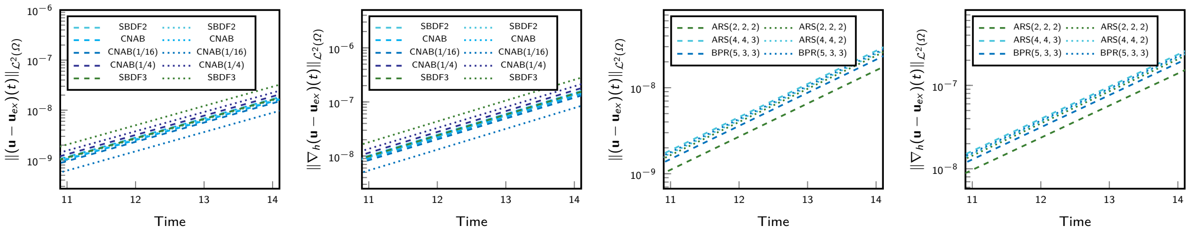


Figure 5.11.: The results from Figure 5.9 and Figure 5.10 zoomed to the time interval $[10.9, 14.1]$.

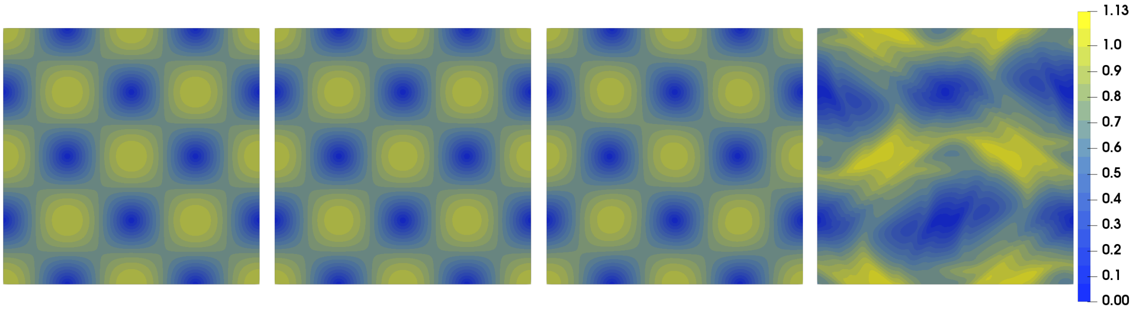


Figure 5.12.: Velocity magnitude of the FEM solution at $t = 10, 20, 30, 40$ computed on the fine mesh with BDM8 finite elements.

5.2.3. Laminar Flow Past a Cylinder

As a second problem we will consider the benchmark problem from the DFG Priority Research Program "Flow Simulations on High Performance Computers" which is formulated in [STD$^+$96] and denoted by the authors as "2D-3". Here the laminar flow around a cylinder inside a channel is considered over a fixed period of time with a time dependent inflow profile.

The domain Ω is a rectangle without a circular obstacle close to the vertical center of the channel

$$\Omega = [0, 2.2] \times [0, 0.41] \backslash \{\mathbf{x} \mid \|\mathbf{x} - (0.2, 0.2)\|_2 \leq 0.05\}.$$

The boundary is composed of three sections. The inflow boundary $\Gamma_{\text{in}} = \{\mathbf{x} \in \Omega \mid x_1 = 0\}$, the outflow boundary $\Gamma_{\text{out}} = \{\mathbf{x} \in \Omega \mid x_1 = 2.2\}$ and the solid wall boundary $\Gamma_{\text{wall}} = \partial\Omega \backslash \Gamma_{\text{in}} \cup \Gamma_{\text{out}}$. On the inflow boundary we prescribe a time-dependent parabolic inflow profile with Dirichlet boundary conditions. This is given by

$$\mathbf{u}(0, x_2; t) = \sin(\frac{\pi t}{8}) \cdot 4 \left(\frac{3}{2} \cdot \overline{\mathbf{u}} \right) \cdot \frac{x_2(0.41 - x_2)}{0.41^2}$$

with a mean velocity $\overline{\mathbf{u}}$. On the outflow boundary we prescribe homogeneous Neumann boundary conditions $(-\nu \nabla \mathbf{u} + pI) \cdot \mathbf{n} = 0$ and no-slip Dirichlet boundary conditions on the remaining wall boundary Γ_{wall}. The viscosity is fixed as $\nu = 10^{-3}$. With the reference length $L = 0.1$ being the diameter of the obstacle and the reference velocity $V = 1$ being the maximum of the mean velocity on the inflow profile we get the Reynolds number $Re = 100$.

To discretise the domain we use four triangulations and BDM8, $\boldsymbol{\mathcal{H}}(\text{div})$-conforming, pointwise divergence free finite elements. The elements on the circular obstacle are then curved to give a piecewise 8th order approximation of the circle. The coarsest mesh resolves the circular obstacle with $h_{max}^{cir} = 0.03$ and $h_{max}^{vol} = 0.08$ in the volume while on the finest mesh we have

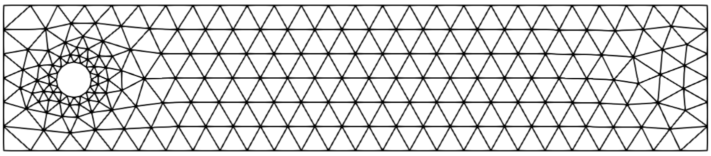


Figure 5.13.: Coarsest mesh for the computation of the reference values of the benchmark problem 2D-3.

$h^{cir}_{max} = 0.001$ and $h^{vol}_{max} = 0.02$. The coarsest of these meshes can be seen in Figure 5.13. The upwinding parameter is again $\gamma = 1$ while the symmetric interior penalty parameter is now chosen as $\sigma = 4k^2$.

For the temporal discretisation we will not consider the full set of IMEX schemes. We have seen in Section 5.2.2 that the temporal error can dominate for the SBDF1 scheme even for a temporally simple problem. We shall therefore not consider this any further. Within the IMEX multistep schemes of expected order two we will only consider the SBDF2 scheme and the CNAB(1/16) method. This is because the SBDF2 scheme is conceptually simple and involves no explicit evaluation of the Stokes part while the CNAB(1/16) gave the marginally best results with large T for the planar vortex problem. Additionally, we will still consider the SBDF3 method in order to check whether it presents any advantages when dealing with temporally more challenging problems compared to the planar vortex problem. Within the IMEX Runge-Kutta regime we will use all three methods considered in the previous example. For each discretisation we consider a large time step close to the stability limit and one small time step.

For this benchmark problem there are five quantities of interest. We consider the forces acting on the circular obstacle, given by the drag and lift coefficients defined as

$$c_D(t) = \frac{2}{L\bar{\mathbf{u}}(t)^2}\int_{\Gamma_{\text{circ}}}\left(\nu\frac{\partial\mathbf{u}(t)}{\partial\mathbf{n}} + p\mathbf{n}\right)\cdot e_x \,\mathrm{d}s$$
$$c_L(t) = \frac{2}{L\bar{\mathbf{u}}(t)^2}\int_{\Gamma_{\text{circ}}}\left(\nu\frac{\partial\mathbf{u}(t)}{\partial\mathbf{n}} + p\mathbf{n}\right)\cdot e_y \,\mathrm{d}s$$

where $L = 0.1$ is the diameter of the obstacle, Γ_{circ} is the boundary of the circle and e_x, e_y are the unit vectors in the x- and y-direction respectively. The reference values are then the maximal drag and lift coefficients over time as well as the time at which they are attained. Furthermore, we compute the pressure difference of the solution at the front and back of the disk

$$\Delta p(t) = p(0.15, 0.2; t) - p(0.25, 0.2; t).$$

The reference value is then $\Delta p(8)$.

We compare our results to those obtained by John and Rang [JR10] using adaptive time-stepping and $\mathbb{Q}_2/\mathbb{P}_1^{\text{disc}}$ $\mathcal{H}^1$-conforming finite elements. Unlike [JR10] we compute the boundary integrals for the reference values directly rather than using a volume term which is equivalent to the boundary term in the $\boldsymbol{\mathcal{H}}^1$-conforming case. This is possible since such boundary integrals are necessary for dG- and $\boldsymbol{\mathcal{H}}(\text{div})$-methods and therefore implemented in `NGSolve`.

The curves for the quantities of interest resulting from the SBDF2 scheme on the coarsest mesh are shown in Figure 5.14 and the velocity solution computed on the same mesh using the SBDF3 method is shown in Figure 5.15. Qualitatively, these results match the reference results, c.f. [Joh16, Example D.9].

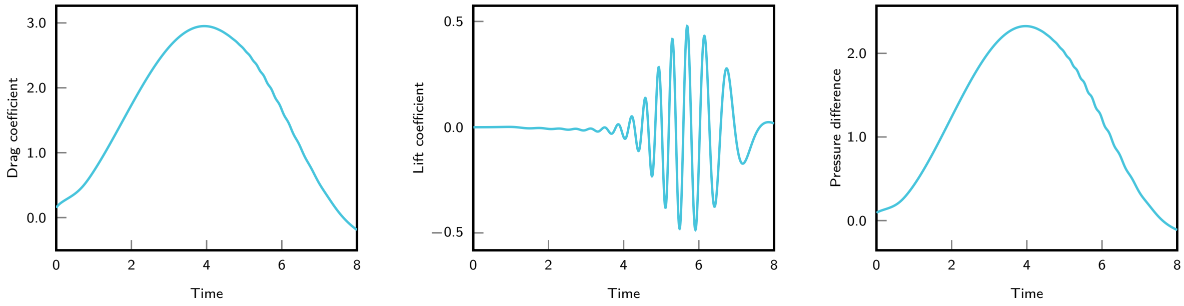


Figure 5.14.: Curves of the quantities of interest for the benchmark problem 2D-3. Results from the coarsest mesh with BDM8 elements and largest time step using the SBDF2 scheme.

The results from our computations are shown in Table 5.9. We see that the BPR(5,3,3) method does not lead to satisfactory results and both mesh and time step refinement gives no better results with this scheme. For all other schemes we observe that we obtain very similar results compared to the reference results. We also see that the spatial resolution is the major contributing factor concerning the accuracy of the results. The results on the finest mesh agree with the reference results up to order 10^{-5} in $t_{D,\max}$ and $c_{D,\max}$, order 10^{-4} in $t_{L,\max}$, order 10^{-5} in $c_{L,\max}$ and order 10^{-7} in $\Delta p(8)$.

Remark 5.8 (Upwinding). We have conducted all numerical experiments shown in this and the previous sections with the upwinding term included by setting $\gamma = 1$ in the dG convective term. Our experience in this thesis has shown that this term is necessary for the planar vortex problem together with large time steps but not relevant for smaller time steps. We observed that the $\mathcal{L}^2$-error was marginally smaller without upwinding for small times t, c.f. Figure 5.16. For the 2d-3 benchmarking problem we did not observe the necessity of the upwinding term even for large time steps.

5.3. Numerical Examples in Three Dimensions

We will now consider numerical examples in 3D to see how the time-integration schemes perform in this situation. Here the CFL bound on the time step and the memory efficiency of the schemes becomes more important as in three dimensions an internal degree of freedom usually has more neighbouring degrees of freedom compared with the situation in two dimensions [Joh16]. Therefore the linear systems are less sparse, requiring more memory and it usually takes longer to solve these systems. This makes larger time steps even more important.

As we have seen above, the SBDF2 method allows on the one hand the largest time steps and on the other hand gives results as accurate as any the other second order IMEX multistep scheme. It is therefore the only second order IMEX multistep scheme we will consider here. Furthermore, we will consider the SBDF3 scheme as we have not yet seen any difference between this and the second order schemes. For the IMEX Runge-Kutta methods we have seen that the BPR(5,3,3) method can lead to unsatisfactory results. Hence, we will only continue to consider the ARS IMEX Runge-Kutta methods.

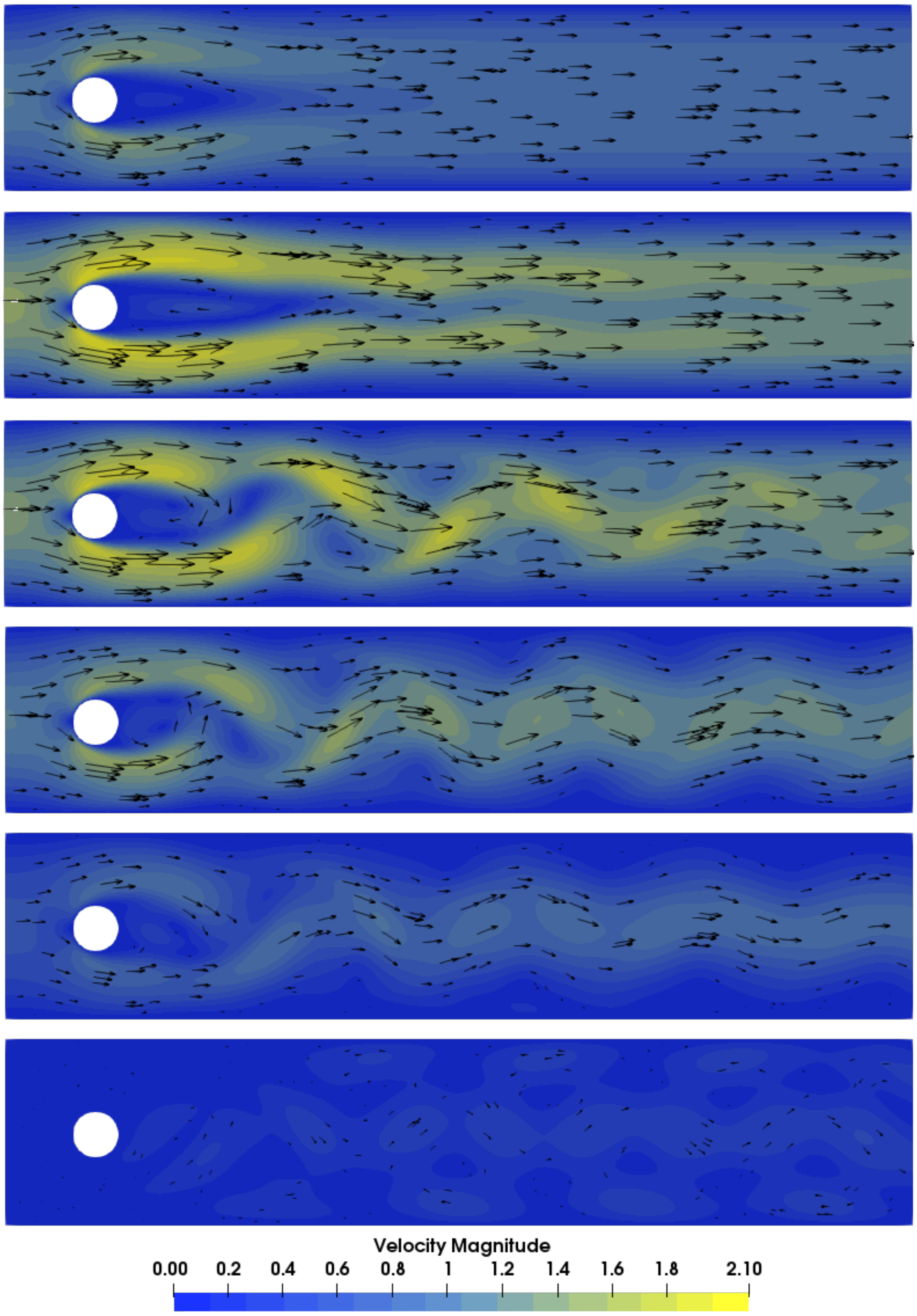


Figure 5.15.: The velocity solution of the problem 2D-3 at time $t = 2, 4, 5, 6, 7$ and 8. Computed on the coarsest mesh with BDM8 elements using the SBDF3 time-integration and the time step $\tau = {}^{1}/_{3000}$.

Method	$h_{max}^{vol}/h_{max}^{cir}$	τ	$t_{D,\max}$	$c_{D,\max}$	$t_{L,\max}$	$c_{L,\max}$	$\Delta p(8)$
SBDF2	0.08/0.03	$1/2000$	3.9365	2.9509305244	5.691	0.4780164260	-0.1114870980
	0.08/0.03	$1/8000$	3.9365	2.9509294999	5.69125	0.4778435307	-0.1115040793
	0.04/0.01	$1/3000$	3.93633	2.9504925238	5.693	0.4778986650	-0.1116266081
	0.04/0.01	$1/8000$	3.93625	2.9504921191	5.693	0.4778292496	-0.1116325692
	0.03/0.005	$1/4000$	3.93625	2.9506955148	5.69275	0.4779168548	-0.1116175367
	0.03/0.005	$1/10000$	3.9363	2.9506952931	5.692875	0.4778827308	-0.1115956342
	0.02/0.001	$1/10000$	3.9363	2.9509039205	5.6928	0.4778977300	-0.1116160903
CNAB($1/16$)	0.08/0.03	$1/2500$	3.9364	2.950929846	5.6912	0.4779023524	-0.1114980666
	0.08/0.03	$1/8000$	3.9365	2.9509294730	5.69125	0.4778389607	-0.1115045057
	0.04/0.01	$1/4500$	3.93622	2.9504921772	5.69288	0.4778390898	-0.1116316386
	0.04/0.01	$1/8000$	3.93625	2.9504920928	5.693	0.4778247174	-0.1116329385
	0.03/0.005	$1/5500$	3.936364	2.9506953349	5.692909	0.4778856147	-0.1116201964
	0.03/0.005	$1/10000$	3.9363	2.950695276	5.6929	0.4778756872	-0.1116210979
	0.02/0.001	$1/10000$	3.9363	2.9509039028	5.6928	0.4778948273	-0.1116163315
SBDF3	0.08/0.03	$1/3000$	3.93633	2.9509294191	5.69133	0.4778310846	-0.1115052133
	0.08/0.03	$1/8000$	3.9365	2.9509294323	5.691375	0.4778320619	-0.1115052005
	0.04/0.01	$1/5500$	3.93636	2.9504920504	5.693091	0.4778172798	-0.1116335411
	0.04/0.01	$1/8000$	3.93625	2.9504920531	5.693	0.4778178578	-0.1116335398
	0.03/0.005	$1/7000$	3.936286	2.95069525	5.692857	0.4778713176	-0.1115930670
	0.03/0.005	$1/10000$	3.9363	2.950695251	5.6929	0.4778713294	-0.1116214888
	0.02/0.001	$1/10000$	3.9363	2.9509038756	5.6928	0.4778904428	-0.1116167246
ARS(2,2,2)	0.08/0.03	$1/1300$	3.936154	2.9509426433	5.690769	0.4775244248	-0.1115253178
	0.08/0.03	$1/5300$	3.936415	2.9509339066	5.691132	0.4777266511	-0.1115119203
	0.04/0.01	$1/2000$	3.936	2.9504994611	5.6925	0.4775783113	-0.1116498784
	0.04/0.01	$1/5300$	3.936226	2.9504949193	5.69283	0.4777117097	-0.1116406309
	0.03/0.005	$1/2600$	3.936154	2.9507009303	5.692692	0.4776776831	-0.1116345357
	0.03/0.005	$1/6600$	3.936212	2.9506975644	5.692727	0.4777867194	-0.1116270960
	0.02/0.001	$1/6600$	3.936212	2.9509060637	5.692727	0.4778049750	-0.1116223963
ARS(4,4,3)	0.08/0.03	$1/1300$	3.936154	2.9509371677	5.690769	0.4775978114	-0.1115221105
	0.08/0.03	$1/3600$	3.936111	2.9509333853	5.691389	0.4777423376	-0.1115111325
	0.04/0.01	$1/2500$	3.936	2.9504953213	5.6928	0.4776907140	-0.1116424726
	0.04/0.01	$1/3600$	3.935833	2.9504944000	5.693056	0.4777270690	-0.1116397483
	0.03/0.005	$1/3200$	3.93625	2.9506978462	5.692813	0.4777729232	-0.1116282958
	0.03/0.005	$1/4450$	3.936264	2.9506971423	5.692747	0.4778016551	-0.1116262324
	0.02/0.001	$1/4450$	3.936264	2.9509056346	5.692747	0.4778199441	-0.1116215338
BPR(5,3,3)	0.08/0.03	$1/1150$	4.071305	2.9515422801	5.690435	0.4773333460	0.0503485002
	0.08/0.03	$1/3000$	4.069667	2.9514740971	5.691	0.4774839120	0.0495976333
	0.04/0.01	$1/2100$	4.070952	2.9510836483	5.692381	0.4774421594	0.0496965050
	0.04/0.01	$1/3000$	4.070667	2.9510613594	5.692667	0.4774834432	0.0494515015
	0.03/0.005	$1/2650$	4.070943	2.9512771026	5.692453	0.4775250147	0.0495495227
	0.03/0.005	$1/3800$	4.070526	2.9512576411	5.692632	0.4775570657	0.0493244109
Ref. [JR10]		$[5 \cdot 10^{-4}, 0.1]$	3.93625	2.950918381	5.6925	0.47787543	-0.11161567

Table 5.9.: Results for the reference values of the benchmark problem "Flow Around a Cylinder" 2D-3 using IMEX multistep methods and IMEX Runge-Kutta methods together with BDM8 elements.

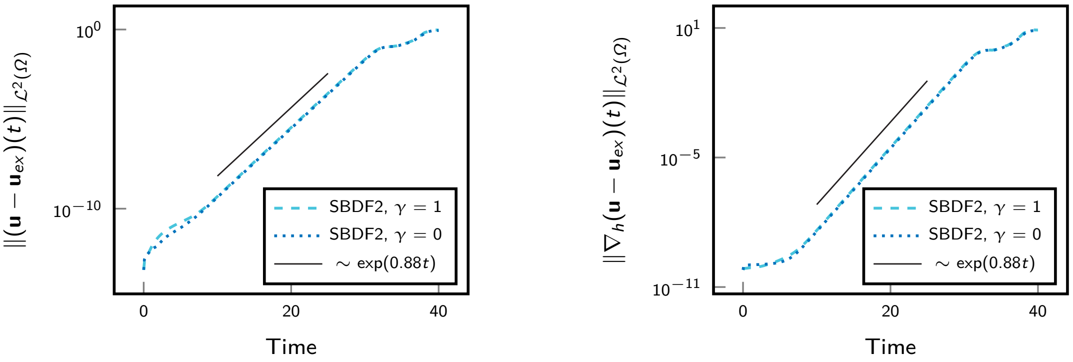


Figure 5.16.: Error development of the planar vortex computed on the fine mesh using BDM8 elements and a time step $\tau = 10^{-4}$ for the SBDF2 scheme with and without upwinding.

5.3.1. General Set-up

The dG-parameters are taken to be $\gamma = 1$ and $\sigma = 4k^2$ as before in the computations in Section 5.2.3. The initial mesh of the cube consists of 6 tetrahedra (this is the coarsest tetrahedral decomposition possible). To obtain finer meshes we take the previous mesh and refine each tetrahedron into 8 tetrahedra. As a result all meshes considered here are structured.

5.3.2. Ethier-Steinman Problem

We consider the analytical solution to the Navier-Stokes equations derived by Ethier and Steinman in [ES94]. For $\Omega = [-1,1]^3$, the initial state is given by

$$\mathbf{u}_0(x,y,z) = -a \begin{pmatrix} e^{ax}\sin(ay \pm dz) + e^{az}\cos(ax \pm dy) \\ e^{ay}\sin(az \pm dx) + e^{ax}\cos(ay \pm dz) \\ e^{az}\sin(ax \pm dy) + e^{ay}\cos(az \pm dx) \end{pmatrix}$$

and

$$\begin{aligned} p_0(x,y,z) = -\frac{a^2}{2}\Big[e^{2ax} + e^{2ay} + e^{2az} &+ 2\sin(ax \pm dy)\cos(az \pm dx)e^{a(y+z)} \\ &+ 2\sin(ay \pm dz)\cos(ax \pm dy)e^{a(z+x)} \\ &+ 2\sin(az \pm dx)\cos(ay \pm dz)e^{a(x+y)}\Big] \end{aligned}$$

for real coefficients a and d. Like the planar vortices this solution was built such that the convective term balances with the pressure gradient and in the Navier-Stokes context the viscous term is balanced by the temporally unsteady term. For $\nu > 0$, the solution is then given by

$$\mathbf{u}(\mathbf{x},t) = \mathbf{u}_0(\mathbf{x})e^{-\nu d^2 t} \quad \text{and} \quad p(\mathbf{x},t) = p_0(\mathbf{x})e^{-2\nu d^2 t}. \tag{5.7}$$

As in [ES94; DE11], we take $a = \pi/4$ and $d = \pi/2$ in (5.7) and the resulting velocity field can be seen in Figure 5.17. We choose the kinematic viscosity to be $\nu = 0.01$, and on the boundary we impose Dirichlet boundary conditions on all 6 faces of the cube according to the analytical solution.

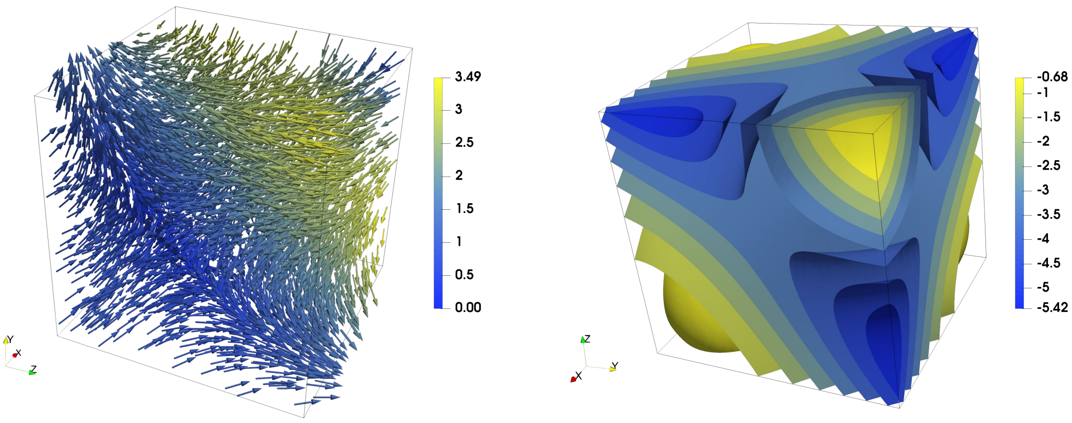


Figure 5.17.: Velocity field and pressure iso-surfaces of the Ethier-Steinman solution at $t = 0$ with $a = {}^{\pi}/_{4}$ and $d = {}^{\pi}/_{2}$.

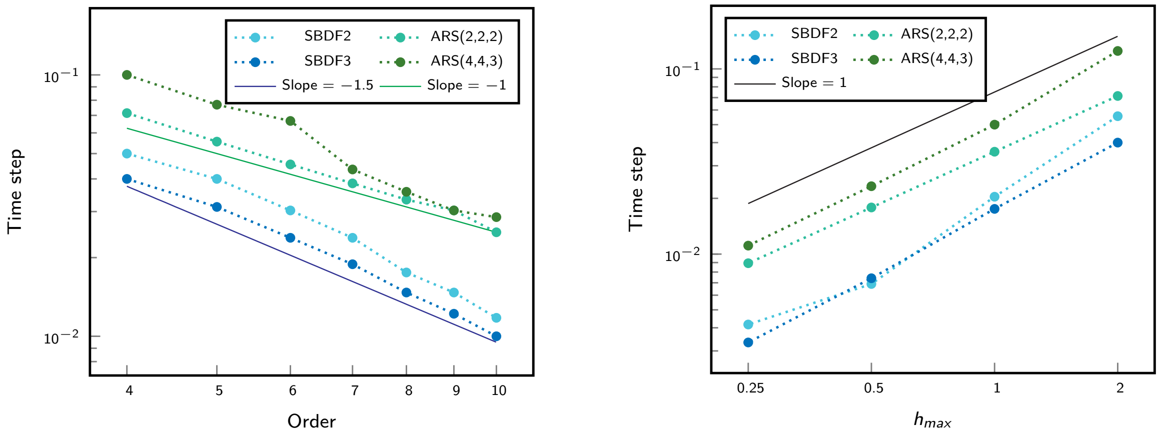


Figure 5.18.: The largest stable time step for the Ethier-Steinman problem. Computed using BDM elements of order k on a mesh consisting of 6 tetrahedra (right) and using on BDM4 elements (left) respectively.

Stability. We investigate if we observe different behaviour in 3D with respect to the CFL condition compared to our computations in 2D, i.e. whether the scaling with respect to k and h_{max} is as before. For this we consider the Ethier-Steinman problem (5.7) once on the initial mesh with BDM elements of order $4 \leq k \leq 10$ and once on a series of uniform refined meshes with BDM4 elements.

The results can be seen in Figure 5.18. We observe that both SBDF schemes scale with respect to the polynomial order k with the factor $3/2$ as in the 2D case. However, for the ARS(2,2,2) we observed a linear scaling with respect to the polynomial order k. It is unclear why the CFL scaling of the IMEX Runge-Kutta schemes is weaker here than in the earlier, convection dominated two dimensional case, and the heuristic information from the analysis above does not help to explain this.

Due to the different scalings with respect to k the ARS(2,2,2) allowed for time steps between 1.4 (at $k = 4$) and 2 (at $k = 10$) times larger than for SBDF2 while the ARS(4,4,3) method allowed for time steps between 2 and 2.5 times larger than SBDF2.

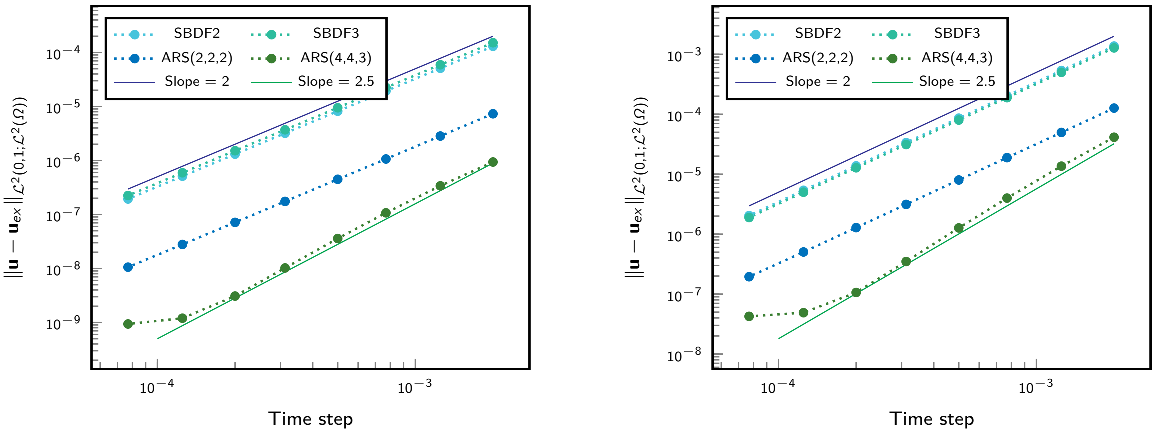


Figure 5.19.: The full $\mathcal{L}^2(0,1;\mathcal{L}^2(\Omega))$-norm (left) and $\mathcal{L}^2(0,1;\mathcal{H}^1(\Omega))$-norm (right) velocity error for the modified Ethier-Steinman problem. Computed using BDM8 elements on a unstructured tetrahedral mesh with $h_{max} = 0.8$ over a series of time steps.

5.3.3. Modified Ethier-Steinman Problem

The Ethier-Steinman benchmarking solution (5.7) has a mild exponential factor in time similar to the planar vortex problem in Section 5.2. To introduce larger variations in time making the time-integration more difficult we modify the Ethier-Steinman solution with respect to time so as to preserve the initial condition and the balance of the convective term with the pressure gradient by multiplying the velocity with a cosine function in time. The aim is to see the temporal order of the considered IMEX schemes which we were unable to observe before.

The new analytical solution reads

$$\mathbf{u}(\mathbf{x},t) = \mathbf{u}_0(\mathbf{x})\cos(2\pi t)e^{-\nu d^2 t} \quad \text{and} \quad p(\mathbf{x},t) = p_0(\mathbf{x})\cos^2(2\pi t)e^{-2\nu d^2 t}. \tag{5.8}$$

Clearly the time derivative can no longer be balanced by the diffusion term. Using the product rule we see

$$\begin{aligned}\partial_t \mathbf{u}(\mathbf{x},t) &= \mathbf{u}_0(\mathbf{x})(-\nu d^2\cos(2\pi t)e^{-\nu d^2 t} - 2\pi\sin(2\pi t)e^{-\nu d^2})\\ &= -\nu\Delta\mathbf{u} - \mathbf{u}_0(\mathbf{x})2\pi\sin(2\pi t)e^{-\nu d^2 t}.\end{aligned}$$

So to fulfil the incompressible Navier-Stokes equations we have to introduce a body forcing term

$$\mathbf{f}(\mathbf{x},t) = -\mathbf{u}_0(\mathbf{x})2\pi\sin(2\pi t)e^{-\nu d^2 t}.$$

We shall refer to this problem as the *modified Ether-Steinman* problem.

Temporal Convergence. We consider the modified Ethier-Steinman problem over the time interval $[0,1]$. We take $a = \pi/4$, $d = \pi/4$ and $\nu = 0.01$ in (5.8) as before. On the boundary we again impose Dirichlet boundary conditions according to the exact solution. For this temporal convergence study we use BDM8 elements on an unstructured tetrahedral mesh with $h_{max} = 0.8$ containing 169 tetrahedra. The time steps are chosen between $1/500$ and $1/13000$ for all four schemes considered here so as to see temporal convergence of the different methods.

The results are shown in Figure 5.19. Unlike the planar vortex example in Subsection 5.2.2 we observe that the SBDF2 and ARS(2,2,2) schemes are of second order as expected.

Similarly [BPR13] observed second order convergence for the ARS(2,2,2) method applied to a hyperbolic system of conservation laws with diffusive relaxation. However for the ARS(4,4,3) scheme we only observe convergence of order $5/2$. This compares with [BPR13], where an order of convergence between 2.4 and 3.3 was observed for the ARS(4,4,3) method. The SBDF3 only converges with quadratically with the time step. In fact, the SBDF3 method gives virtually identical results as SBDF2. It is not clear why we observe this drop in the order of the method since the k-step BDF schemes applied to DAE systems (fully implicit) are of order k for $k \leq 6$ [HW96; KM06].

With respect of accuracy we see that the ARS(2,2,2) method gives results which are more accurate by an order of magnitude compared to SBDF2. Furthermore, the ARS(4,4,3) method gives results which are up to an order of magnitude more accurate than the ARS(2,2,2) results. To take into account the additional effort, we compare the results from the ARS(2,2,2) method with a time step twice as large as for SBDF2. We find that the ARS(2,2,2) results are more accurate by a factor of 4.5 in the $\mathcal{L}^2(\mathcal{L}^2)$-error and a factor of 2 in the $\mathcal{L}^2(\mathcal{H}^1)$-error. However, the ARS(4,4,3) results are more accurate by a factor between 8 and 21 and between 2 and 8 with respect to the $\mathcal{L}^2(\mathcal{L}^2)$ and $\mathcal{L}^2(\mathcal{H}^1)$ errors compared to SBDF2 when considering the same computational effort, i.e. time steps four times as large.

Remark 5.9 (Upwinding). As in the 2d case all results presented here were computed with the upwinding term in the discretised convection operator included with $\gamma = 1$. In practice we observed for the 3d computations considered here that the upwinding term was necessary in the CFL stability limit, however, with a time step 1.1 times below this there was no significant difference between the results with and without the upwinding term.

5.4. Assessment of the Schemes

We summarise the results obtained in Sections 5.2 and 5.3:

SBDF1

- Allows the largest time steps in the CFL limit for all multistep methods on coarse meshes and for $2 \leq k \leq 10$.
- CFL scaling $\tau \leq C(h/k)^{1.5}$ better with respect to k but worse with respect to h than expected.
- The temporal error can dominate for problems with simple behaviour in time.
- First order convergence in time of the scheme was observed.

Second order multistep schemes

- We did not observe any significant difference between the schemes concerning accuracy.
- CFL bound $\tau \leq Ch/k^{1.5}$ observed in both the convection dominated 2d case and diffusion dominated 3d case, i.e. scaling with respect to h as expected and better with respect to k than expected.
- The SBDF2 scheme allows the largest time step in the CFL limit.
- Second order convergence with respect to the time step was observed for the SBDF2 scheme for the modified Ethier-Steinman problem.

SBDF3

- With respect to accuracy, the SBDF3 method performed similarly to the SBDF2 method throughout.
- Same CFL scaling as for second order schemes but smaller constant than SBDF2, leading to smaller time steps in the CFL limit.
- Order reduction with respect to the time step observed so that we only obtained second order convergence with the additional effort of computing the matrix and preconditioner for the SBDF3 steps and having to use smaller time steps compared with SBDF2.

ARS(2,2,2)

- Second order convergence in time was observed.
- Since the first step is second order this method can give slightly more accurate results for the same effort compared with SBDF2.
- The CFL restriction results in a higher computational cost compared with the SBDF2 method in the stability limit.
- CFL condition $\tau \leq Ch/k^{1.5}$ observed in convection dominated case but $\tau \leq Ch/k$ observed in the diffusion dominated case.
- More accurate than SBDF2 by a factor 4.5 in the $\mathcal{L}^2$-error and a factor 2 in the $\mathcal{H}^1$-error for the same computational cost observed for the modified Ethier-Steinman problem.

ARS(4,4,3)

- Order of $^5/_2$ observed in practice.
- Results more accurate than SBDF2 by an order of magnitude for the same computational cost observed for temporally challenging modified Ethier-Steinman problem.
- Higher computational cost than SBDF3 in the CFL stability limit.

BPR(5,3,3)

- Results for the 2d-3 benchmarking problem where inaccurate and did not converge towards reference results.

We can therefore discard a number of methods. Due to the order reduction of the SBDF3 scheme and the lack of accuracy of the BPR(5,3,3) method we can reject using either of these methods for divergence-free Navier-Stokes computations.

For problems which present simple behaviour in time (such as a mild exponential term) the temporal error dominated in the case of the SBDF1 scheme when a high spatial resolution was considered. This scheme should therefore not be used.

Between the second order IMEX multistep schemes we did not observe any significant differences between the individual schemes with respect to accuracy. In the non-incremental form the CNAB(δ) schemes require us to store the Stokes matrix A in addition to the convection operator C and the system matrix M^*. These schemes therefore require more memory compared with SBDF2 in a non-incremental implementation. These schemes also require smaller time steps in the CFL limit compared with SBDF2. We therefore come to the conclusion that the SBDF2 scheme is the best scheme in the regime of second order multistep methods.

Computationally it is the cheapest method and conceptually very simple, combining the BDF2 method for the Stokes part with a second order extrapolation for the convection term.

In a straight forward implementation of the methods, as we considered in this Chapter, IMEX Runge-Kutta methods also present advantages in the following situations. In temporally challenging problems we have seen that IMEX Runge-Kutta methods give more accurate results for the same computational cost. Even though the ARS(4,4,3) method is not of 3rd order in practice, we have seen that it gives results which are an order of magnitude more accurate than SBDF2 for the same computational effort. Even ARS(2,2,2) presented slightly better results than SBDF2 for the modified Ethier-Steinman problem. We attribute this to the first time step which in the case of IMEX Runge-Kutta schemes is a higher order approximation rather than a first order approximation used to compute the first time step in the second order IMEX multistep methods. This also means that IMEX Runge-Kutta methods use less memory compared with IMEX multistep schemes if the multistep scheme is not carefully implemented. As a result the code for an IMEX Runge-Kutta implementation can be structured more easily in a simple implementation. However, in situations simpler with regard to time the need to solve multiple linear systems per time step in combination with the CFL condition makes IMEX Runge-Kutta methods require computational effort in the CFL limit compared with SBDF2 for the same level of accuracy.

As a result we conclude that each of the SBDF2, ARS(2,2,2) and ARS(4,3,3) methods can be an effective method depending on the given problem. It is therefore advantageous to test each of these schemes and then to take a higher order scheme if significant differences can be observed in the results.

Remark 5.10. The theoretical considerations we have used to study IMEX multistep and IMEX Runge-Kutta methods for the Navier-Stokes equations are limited. In Sections 3.2–3.3 and 4.2 we used a scalar test-problem to obtain stability in the sense that the eigenvalue of the scalar problem has to be contained in the stability domain of the time stepping method. The generalisation of this then takes the eigenvalue of the scalar problem and treats it as the maximal eigenvalue of the multidimensional problem. This approach assumes that all the operators are linear to have the concept of eigenvalues but the convection operator we treat explicitly is non-linear. Consequently, this only shows which schemes should have the best constant in the time step restriction. The CFL condition for stability in Section 5.1 brought us closer to a practical time step restriction. In practice we observed that this bound seems sharp with respect to h as already noted by [JL04; KKW17]. Furthermore, we also observed that this bound is not sharp with respect to the polynomial order k, a result previously observed by [KKW17]. However, the CFL bound was only shown for a scalar transport problem rather than for the full Navier-Stokes problem.

Essentially the full discrete analysis of the IMEX multistep and IMEX Runge-Kutta schemes applied to the Navier-Stokes equations is missing. For the first order SBDF1 (or implicit/explicit Euler) method applied to the incompressible Navier-Stokes equations a fully discrete analysis has been conducted in [He08]. For the second order CNAB method fully discrete analysis has been conducted in [HS07; Ton04; MT98]. Unfortunately, the stability analysis in these papers assumes that the time step is bounded by negative powers of the viscosity and ignores any dependence on the polynomial order k of the FEM space.

In our literature research we did not find any fully discrete analysis for IMEX Runge-Kutta methods for the Navier-Stokes equations. This is therefore an open problem. Another open problem is a discrete analysis giving a sharp stability bound on the time step to give the stability bounds observed in practice a theoretical foundation.

Remark 5.11 (Other time-stepping techniques for the Navier-Stokes equations). There is a wide range of alternative time-integration approaches for the incompressible Navier-Stokes equations besides the additive IMEX splittings we have considered in this thesis. Within an IMEX regime, i.e. treating the Stokes part implicitly and the convection part explicitly, multiplicative decomposition methods such as the operator-integrator-factor splitting introduced by [MPR90] and used in [LS16] are possible. This method presents the advantage over additive IMEX schemes that a different time step can be used for the implicit and explicit part since the two part are solved in separate steps. This circumvents the CFL condition of the explicit part restricting the entire method which was the case for our schemes.

Other methods include projection methods such as pressure-correction methods. An overview of projection methods for the Navier-Stokes equations can be found in [GMS06]. Here the the velocity and pressure are advanced in separate sub-steps such that the saddle-point structure of the coupled system is avoided. A method which combines the velocity-pressure splitting approach and the IMEX approach to to diffusion and convection was presented in [KIO91]. We note however that such methods which split the velocity and pressure entirely should not be used in conjunction with pointwise divergence FEM, as presented in Chapter 2, as it is the coupling between the velocity and pressure is the essential component which enforces the divergence constraint pointwise, c.f. Lemma 2.18.

A different type of method which are referred to as IMEX methods by [Joh16] are schemes which do not treat the entire convective term explicitly. Instead the convective term in linearised by substituting the advection velocity field with some extrapolation term consisting of previous velocities.

Fully implicit methods are also possible. In this case Newton's method or a Picard iteration can be used to solve the non-linear system arising in each time step, c.f. for example [Joh16] for a comparison of these two approaches.

Related to Runge-Kutta methods are Rosenbrock-type methods [HW96]. The are essentially Newton linearisations of DIRK methods, where the Jacobian is only updated once per time step rather than at every intermediate stage.

6. Conclusion and Outlook

In this thesis we considered implicit-explicit time splitting schemes for the temporal discretisation of the incompressible Navier-Stokes equations in connection with $\boldsymbol{\mathcal{H}}(\mathrm{div})$-conforming exactly divergence free FEM. In this chapter we summarise the most important aspects focused on in this thesis, cover the results attained and explain some of the remaining problems.

Summary

The Navier-Stokes equations. We began in Chapter 1 by introducing the incompressible Navier-Stokes equations and covering notation and results from Functional Analysis such as weak derivatives and the Ladyzhenskaya-Babuška-Brezzi condition in an abstract setting. We then formulated the weak form of the Navier-Stokes with both time-independent and time-dependent test functions and covered the equivalence of these two formulations. After this we cited the solvability of the Navier-Stokes equations in the weak formulation and remarked on the open problems connected with this.

Spatial semi-discretisation. In Chapter 2 we covered the spatial semi-discretisation of the incompressible Navier-Stokes equations using $\boldsymbol{\mathcal{H}}(\mathrm{div})$-conforming finite elements. We showed that a function from the broken space $\boldsymbol{\mathcal{H}}(\mathrm{div}; \mathcal{T}_h)$ is also in the global space $\boldsymbol{\mathcal{H}}(\mathrm{div}; \Omega)$ if and only if the function is continuous in the normal component across element facets. As a result we required dG formulations of the multi-linear forms in the weak formulation of the Navier-Stokes equations. We therefore derived the Symmetric Interior Penalty method for the Poisson problem, as a prototype for the diffusion term, showed that in the $\boldsymbol{\mathcal{H}}(\mathrm{div})$-conforming context we can use the continuous velocity-pressure coupling term in the discrete setting and we derived a dG formulation of the convective term with an optional upwind stabilisation term.

Having formulated the spatially semi-discrete Navier-Stokes problem we proved the existence and uniqueness of solutions to this problem for inf-sup stable velocity-pressure space pairs. Following this we cited recently proven error estimates and convergence results which are both pressure- and Reynolds-semi-robust. We concluded the chapter with an overview of possible finite element approximations of $\boldsymbol{\mathcal{H}}(\mathrm{div})$ leading to inf-sup stable and pointwise divergence-free methods.

IMEX Multistep schemes. Chapter 3 was concerned with the temporal discretisation of the spatial semi-discrete Navier-Stokes equations using implicit-explicit multistep methods. Following the cited literature we derived IMEX multistep schemes up to formal order three for a general ODE consisting of two parts. We then analysed these schemes extending the work of the literature cited to the SBDF3 method. Within this analysis we considered the stability region of the explicit part of the schemes resulting from a scalar test problem as is customary in ODE analysis and constructed the restrictions within this region necessary for A-stability of the implicit part of the method. Furthermore, we considered $A(\alpha)$-stability of these schemes, and constructed theoretical and numerical estimates of α.

IMEX Runge-Kutta methods. In Chapter 4 we discussed basic notation and definitions regarding Runge-Kutta methods. We presented stiffly accurate IMEX Runge-Kutta methods which we considered to be suitable for the application of pointwise divergence free FEM and available in the cited literature. Of these we restricted ourself to a subset of schemes which we could expect to be competitive with respect to the computational cost of the schemes. For this remaining set we then computed the respective stability domains to give a qualitative prediction of the relative time step restrictions.

Numerical experiments. In Section 5.1 we proved the CFL condition $\tau \leq C^h/_{k^2}$ by considering a scalar transport problem. Section A then covered some implementational aspects. With Section 5.2 we began the numerical simulations to test the performance of the IMEX schemes coved in Chapter 3 and Chapter 4 in practice. For this we used the finite element package `NGSolve`. In two dimensions we considered the planar vortex problem at $Re = 10^5$ on a periodic square and the Schäfer-Turek benchmark problem 2D-3 which contains a time-dependent inflow boundary condition. In three dimensions we considered the Ethier-Steinman problem as well as a modified version thereof to make the resulting problem challenging with respect to the temporal integration.

For the first problem we observed that the expected CFL condition $\tau \leq C^h/_{k^2}$ was too strong with respect to the polynomial order k and that in fact $\tau \leq C^h/_{k^{3/2}}$ gives a sharp bound on the time step for stability. We also saw that of the multistep schemes considered the SBDF1 scheme allowed for the largest time step with respect to k and the SBDF2 scheme allowed the largest time step within the second order multistep schemes throughout. Furthermore, we showed that in the CFL stability limit IMEX Runge-Kutta methods allow larger time steps than multistep schemes, however, the additional computational effort resulting from having to solve additional linear systems in each time step cannot be offset by using larger time steps. With respect to accuracy we showed that the SBDF1 scheme is a first order method. However, for all higher order method we were unable to see the temporal order of convergence in this example, since the CFL restriction on the time step resulted in the dominance of the spatial discretisation error.

The second problem we computed in two dimensions we consider to be temporally more challenging than the first due to the time dependent inflow profile. However, here we also did not observe any significant difference between the results from the schemes used except for the BPR(5,3,3) method wich yielded incorrect results. The main factor influencing the accuracy of the results was the spatial discretisation.

With the Ethier-Steinman problem in three dimensions we looked again at the CFL stability limit in a diffusion dominated situation. Here we saw a milder CFL restriction for the ARS(2,2,2) method than before with the scaling $\tau \leq C^h/_k$. Nevertheless, the CFL scaling with respect to h and k for the IMEX multistep schemes remained as in the two dimensional, convection dominated case.

To establish the temporal order of the schemes expected to be of order greater than one, something we were unable to verify with the previous problems, we constructed a problem with more complicated behaviour in time. We achieved this by multiplying the Ethier-Steinman velocity-solution with $\cos(2\pi t)$ and adapting the pressure and right-hand side such that this new velocity-pressure pair is a solution to the Navier-Stokes equations. Using this problem we showed that the SBDF2 and ARS(2,2,2) methods are of order two as expected. However, we also showed that the SBDF3 scheme is not of third order but only displays quadratic convergence with respect to the time step. Additionally we showed that the ARS(4,4,3) is also not of order three as expected but is of order $^5/_2$. We also observed that the ARS(4,4,3) method can yield results the error of which is an order of magnitude smaller than those given by the SBDF2 method for the same computational effort.

Overall, we concluded in Section 5.4 with an assessment of the schemes presented in this thesis. We concluded that dependant on the specific problem at hand (i.e. if the problem is difficult with respect to the time integration or how memory intensive the problem is and how much time is needed to assemble the matrices) the SBDF2, ARS(2,2,2) or ARS(4,4,3) scheme can be the most effective IMEX method.

Open Problems

With respect to the fully discrete analysis of IMEX methods applied to the incompressible Navier-Stokes equations further work is required. As we have discussed, fully discrete analysis is available for the SBDF1 and CNAB schemes [He08; HS07; MT98; Ton04]. However, these result do not immediately reflect the CFL stability condition observed in this thesis and other works, e.g. [JL04; KKW17]. Also the currently available stability analysis does not account for higher order methods which have been used in conjunction with IMEX methods in this thesis and for example [KKW17; LS16; SJL$^+$18; SLL$^+$18]. With respect to the CFL condition it would be interesting to find a Navier-Stokes flow example where the classical CFL condition $\tau \leq C^h/_{k^2}$ is a sharp bound on the time step or alternatively, give a theoretical foundation for the milder CFL condition $\tau \leq C^h/_{k^{3/2}}$ observed here and in the cited literature. Furthermore, a theoretical explanation to the stronger CFL condition $\tau \leq (^h/_k)^{3/2}$ of the SBDF1 scheme remains an open problem.

We did not find any fully discrete analysis of IMEX Runge-Kutta methods applied to the incompressible Navier-Stokes equations in the literature. This is therefore also a problem which should be considered in further research. Here it would also be relevant to find the reason behind the different CFL scalings of the ARS methods observed between the two and three dimensional examples.

Concerning the order reduction observed for the SBDF3 and ARS(4,4,3) methods, a theoretical understanding is missing. Furthermore, during our numerical simulations we successively stopped considering methods which are more complicated or which gave inaccurate results compared with the other remaining schemes. As a result we did not come to apply these schemes to the temporally difficult problem and thus we did not observe the actual temporal order of schemes such as the CNAB scheme. Investigating the order of the schemes we did not consider in Section 5.3 which could be relevant for other applications also remains open for further work.

Finally, a comparison regarding accuracy and computational effort in practice between the IMEX approach taken here and other methods such as using Newton's method or a Picard iteration is yet to be done. Here it would be of further interest in which situations the different approaches perform best.

Appendix A. Implementational Aspects

A.1. IMEX Multistep Schemes

To avoid the modification of the right-hand side resulting from time-independent non-homogeneous boundary conditions, as described in Section 2.2.1, we bring the schemes into incremental form to homogenise the boundary conditions, i.e. we solve for $\Delta\mathbf{u}^i := \mathbf{u}^{n+1} - \mathbf{u}^n$.

SBDF1 Scheme. The SBDF1 scheme reads

$$\frac{1}{\tau}\Big(M\mathbf{u}^{n+1} - M\mathbf{u}^n\Big) + A(\mathbf{u}^{n+1}, p^{n+1}) + C(\mathbf{u}^n) = 0$$

with the mass-matrix M, the Stokes matrix A and the convection operator C. Taking the implicit terms to the left-hand side and the explicit terms to the right-hand side then yields

$$M\mathbf{u}^{n+1} + \tau A(\mathbf{u}^{n+1}, p^{n+1}) = M\mathbf{u}^n - \tau C(\mathbf{u}^n).$$

Denoting $M^* := M + \tau A$ we therefore get for the increment $\Delta\mathbf{u}^{n+1} := \mathbf{u}^{n+1} - \mathbf{u}^n$

$$M^*(\Delta\mathbf{u}^{n+1}, \Delta p^{n+1}) = -\tau(A(\mathbf{u}^n, p^n) + C(\mathbf{u}^n)).$$

Second order Schemes. General second order IMEX schemes are of the form

$$\begin{aligned}\frac{1}{\tau}\Big(a_{-1}M\mathbf{u}^{n+1} + a_0 M\mathbf{u}^n + a_1 M\mathbf{u}^{n-1}\Big) + b_{-1}A(\mathbf{u}^{n+1}, p^{n+1}) + b_0 A(\mathbf{u}^n, p^n)&\\ + b_1 A(\mathbf{u}^{n-1}, p^{n-1}) + c_0 C(\mathbf{u}^n) + c_1 C(\mathbf{u}^{n-1}) = 0.&\end{aligned}$$

Multiplying this by τ and taking all explicit terms to the right-hand side gives

$$\begin{aligned}a_{-1}M\mathbf{u}^{n+1} + \tau b_{-1}A(\mathbf{u}^{n+1}, p^{n+1}) = -a_0 M\mathbf{u}^n - a_1 M\mathbf{u}^{n-1} - \tau\Big[b_0 A(\mathbf{u}^n, p^n)&\\ + b_1 A(\mathbf{u}^{n-1}, p^{n-1}) + c_0 C(\mathbf{u}^n) + c_1 C(\mathbf{u}^{n-1})\Big].&\end{aligned}$$

Denoting the right-hand side operator by M^* we get

$$\begin{aligned}M^*(\Delta\mathbf{u}^{n+1}, \Delta p^{n+1}) = -(a_{-1} + a_0)M\mathbf{u}^n - a_1 M\mathbf{u}^{n-1}&\\ - \tau\Big((b_{-1} + b_0)A(\mathbf{u}^n, p^n) + b_1 A(\mathbf{u}^{n-1}, p^{n-1}) + c_0 C(\mathbf{u}^n) + c_1 C(\mathbf{u}^{n-1})\Big).&\end{aligned}$$

Inserting the *SBDF2* coefficients $(a_{-1}, a_0, a_1) = (3/2, -2, 1/2)$, $(b_{-1}, b_0, b_1) = (1, 0, 0)$ and $(c_0, c_1) = (2, -1)$ into the above equation then gives

$$M^* = \frac{3}{2}M + \tau A$$

and for the right-hand side

$$\frac{1}{2}M\mathbf{u}^n - \frac{1}{2}M\mathbf{u}^{n-1} - \tau\left(A(\mathbf{u}^n, p^n) + 2C(\mathbf{u}^n) - C(\mathbf{u}^{n-1})\right).$$

The *CNAB* coefficients $(a_{-1}, a_0, a_1) = (1, -1, 0)$, $(b_{-1}, b_0, b_1) = (1/2, 1/2, 0)$ and $(c_0, c_1) = (3/2, -1/2)$ then gives

$$M^* = M + \frac{\tau}{2}A$$

while on the right-hand side we have

$$-\tau\left(A(\mathbf{u}^n, p^n) + \frac{3}{2}C(\mathbf{u}^n) - \frac{1}{2}C(\mathbf{u}^{n-1})\right).$$

The *CNAB(1/16)* coefficients $(a_{-1}, a_0, a_1) = (1, -1, 0)$, $(b_{-1}, b_0, b_1) = (9/16, 3/8, 1/16)$ and $(c_0, c_1) = (3/2, -1/2)$ then gives the matrix for the left-hand side as

$$M^* = M + \tau\frac{9}{16}A$$

and for the right-hand side

$$-\tau\left(\frac{15}{16}A(\mathbf{u}^n, p^n) + \frac{1}{16}A(\mathbf{u}^{n-1}, p^{n-1}) + \frac{3}{2}C(\mathbf{u}^n) - \frac{1}{2}C(\mathbf{u}^{n-1})\right).$$

The *CNAB(1/4)* coefficients $(a_{-1}, a_0, a_1) = (1, -1, 0)$, $(b_{-1}, b_0, b_1) = (3/4, 0, 1/4)$ and $(c_0, c_1) = (3/2, -1/2)$ then give

$$M^* = M + \tau\frac{3}{4}A$$

and for the right-hand side we have

$$-\tau\left(\frac{3}{4}A(\mathbf{u}^n, p^n) + \frac{1}{4}A(\mathbf{u}^{n-1}, p^{n-1}) + \frac{3}{2}C(\mathbf{u}^n) - \frac{1}{2}C(\mathbf{u}^{n-1})\right).$$

SBDF3 Scheme. As with the previous schemes we bring the SBDF3 scheme into incremental form to homogenise the boundary conditions. The resulting scheme is

$$\begin{aligned} M^*(\Delta\mathbf{u}^{n+1}, \Delta p^{n+1}) = \frac{7}{6}M\mathbf{u}^n - \frac{3}{2}M\mathbf{u}^{n-1} + \frac{1}{3}M\mathbf{u}^{n-2} \\ - \tau\left(A(\mathbf{u}^n, p^n) + 3C(\mathbf{u}^n) - 3C(\mathbf{u}^{n-1}) + C(\mathbf{u}^{n-2})\right) \end{aligned}$$

with

$$M^* = \frac{11}{6}M + \tau A.$$

A.2. IMEX Runge-Kutta Schemes

As for the multistep schemes we bring the Runge-Kutta schemes into incremental form to homogenise Dirichlet boundary-conditions. Here the increment is defined as $\Delta\mathbf{u}^i := \tilde{\mathbf{u}}^i - \mathbf{u}^n$ where $\tilde{\mathbf{u}}^i$ is the i-th stage of the Runge-Kutta method. The resulting incremental form of a general IMEX-RK method is described in Algorithm A.1.

```
Data: $\mathbf{u}^n$
Set $\tilde{\mathbf{u}}^1 = \mathbf{u}^n$
for $i = 2$ to $s+1$ do
    Evaluate $C^i = C(\tilde{\mathbf{u}}^i)$
    Evaluate $S^i = A(\tilde{\mathbf{u}}^i)$
    Solve $(M + \tau a_{i+1,i+1} A)\Delta\tilde{\mathbf{u}}^{i+1} = -\tau \sum_{j=1}^{i} \left\{ \widehat{a}_{i+1,j} C^j + a_{i+1,j} S^j \right\} - \tau a_{i+1,i+1} S^1$
    Set $\tilde{\mathbf{u}}^{i+1} = \Delta\tilde{\mathbf{u}}^{i+1} + \mathbf{u}^n$
end
Result: $\mathbf{u}^{n+1} = \tilde{\mathbf{u}}^{s+1}$
```

Algorithm A.1.: Incremental form of IMEX Runge-Kutta scheme using a stiffly accurate method.

A.3. Time dependent Dirichlet Data

For time-dependent Dirichlet boundary conditions we directly use the non-incremental form of both the IMEX multistep and IMEX Runge-Kutta methods. In order to solve for homogeneous Dirichlet conditions we split the velocity

$$\mathbf{u}^{n+1} = \mathbf{u}_0^{n+1} + \mathbf{u}_D^{n+1} \tag{A.1}$$

where $\mathbf{u}_0$ admits homogeneous Dirichlet conditions and $\mathbf{u}_D^{n+1}$ is some known velocity function with the correct boundary conditions. We then solve for $\mathbf{u}_0^{n+1}$ by taking $M^*\mathbf{u}_D^{n+1}$ to the right-hand side, i.e.

$$M^*\mathbf{u}_0^{n+1} = \widetilde{f}(\mathbf{u}^n) - M^*\mathbf{u}_D^{n+1}$$

where $\widetilde{f}(\mathbf{u}^n)$ is the right-hand side resulting from the forcing term and all explicit terms of the time stepping scheme. The new solution then obtained by inserting the two components back into (A.1).

Bibliography

[ARS97] U. M. Ascher, S. J. Ruuth and R. J. Spiteri. Implicit-Explicit Runge-Kutta Methods for Time-dependent Partial Differential Equations. *Applied Numerical Mathematics*, 25(2-3):151–167, November 1997. DOI: 10.1016/s0168-9274(97)00056-1.

[ARW95] U. M. Ascher, S. J. Ruuth and B. T. R. Wetton. Implicit-Explicit Methods for Time-dependent Partial Differential Equations. *SIAM Journal on Numerical Analysis*, 32(3):797–823, June 1995. DOI: 10.1137/0732037.

[BBF13] D. Boffi, F. Brezzi and M. Fortin. *Mixed Finite Element Methods and Applications*. Springer Series in Computational Mathematics. Springer Berlin Heidelberg, 2013. DOI: 10.1007/978-3-642-36519-5.

[Ber88] A. L. Bertozzi. Heteroclinic Orbits and Chaotic Dynamics in Planar Fluid Flows. *SIAM Journal on Mathematical Analysis*, 19(6):1271–1294, November 1988. DOI: 10.1137/0519093.

[BF13] F. Boyer and P. Fabrie. *Mathematical Tools for the Study of the Incompressible Navier-Stokes Equations and Related Models*. Applied Mathematical Sciences. Springer New York, 2013. DOI: 10.1007/978-1-4614-5975-0.

[BIL06] L. C. Berselli, T. Iliescu and W. J. Layton. *Mathematics of Large Eddy Simulation of Turbulent Flows*. Scientific Computation. Springer-Verlag, 2006. ISBN: 3-540-26316-0. DOI: 10.1007/b137408.

[Bos09] S. Boscarino. On an Accurate Third Order Implicit-Explicit Runge-Kutta Method for Stiff Problems. *Applied Numerical Mathematics*, 59(7):1515–1528, July 2009. DOI: 10.1016/j.apnum.2008.10.003.

[BPR13] S. Boscarino, L. Pareschi and G. Russo. Implicit-Explicit Runge–Kutta Schemes for Hyperbolic Systems and Kinetic Equations in the Diffusion Limit. *SIAM Journal on Scientific Computing*, 35(1):A22–A51, January 2013. DOI: 10.1137/110842855.

[BPR17] S. Boscarino, L. Pareschi and G. Russo. A Unified IMEX Runge–Kutta Approach for Hyperbolic Systems with Multiscale Relaxation. *SIAM Journal on Numerical Analysis*, 55(4):2085–2109, January 2017. DOI: 10.1137/m1111449.

[CdFN01] M. P. Calvo, J. de Frutos and J. Novo. Linearly Implicit Runge-Kutta Methods for Advection-reaction-diffusion Equations. *Applied Numerical Mathematics*, 37(4):535–549, June 2001. DOI: 10.1016/s0168-9274(00)00061-1.

[CKS05] B. Cockburn, G. Kanschat and D. Schötzau. A Locally Conservative LDG Method for the Incompressible Navier-Stokes Equations. *Mathematics of Computation*, 74(251):1067–1096, October 2005. DOI: 10.1090/s0025-5718-04-01718-1.

[CKS06] B. Cockburn, G. Kanschat and D. Schötzau. A Note on Discontinuous Galerkin Divergence-free Solutions of the Navier-Stokes Equations. *Journal of Scientific Computing*, 31(1-2):61–73, September 2006. DOI: 10.1007/s10915-006-9107-7.

[DAL15] H. Dallmann, D. Arndt and G. Lube. Local Projection Stabilization for the Oseen Problem. *IMA Journal of Numerical Analysis*, 36(2):796–823, July 2015. DOI: 10.1093/imanum/drv032.

[Dav04] T. A. Davis. Algorithm 832. *ACM Transactions on Mathematical Software*, 30(2):196–199, June 2004. DOI: 10.1145/992200.992206.

[DE11] D. A. Di Pietro and A. Ern. *Mathematical Aspects of Discontinuous Galerkin Methods*. Mathématiques et Applications. Springer Berlin Heidelberg, 2011. DOI: 10.1007/978-3-642-22980-0.

[EG04] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*. Springer New York, 2004. DOI: 10.1007/978-1-4757-4355-5.

[ES94] C. R. Ethier and D. A. Steinman. Exact Fully 3d Navier-Stokes Solutions for Benchmarking. *International Journal for Numerical Methods in Fluids*, 19(5):369–375, September 1994. DOI: `10.1002/fld.1650190502`.

[Eva98] L. C. Evans. *Partial Differential Equations.* Graduate Studies in Mathematics. American Mathematical Society, 1998. ISBN: 978-0-8218-4974-3.

[Fef00] C. L. Fefferman. Official Problem Description: Existence and Smoothness of the Navier-Stokes Equation. 2000. URL: `http://www.claymath.org/sites/default/files/navierstokes.pdf`.

[FHV97] J. Frank, W. Hundsdorfer and J. G. Verwer. On the Stability of Implicit-Explicit Linear Multistep Methods. *Applied Numerical Mathematics*, 25(2-3):193–205, November 1997. DOI: `10.1016/s0168-9274(97)00059-7`.

[Fis14] G. Fischer. *Lineare Algebra.* Springer Fachmedien Wiesbaden, 2014. DOI: `10.1007/978-3-658-03945-5`.

[FWK18] N. Fehn, W. A. Wall and M. Kronbichler. Efficiency of High-performance Discontinuous Galerkin Spectral Element Methods for Under-resolved Turbulent Incompressible Flows. *International Journal for Numerical Methods in Fluids*, May 2018. DOI: `10.1002/fld.4511`.

[Gal00] G. P. Galdi. An Introduction to the Navier-Stokes Initial-boundary Value Problem. In *Fundamental Directions in Mathematical Fluid Mechanics*, pages 1–70. Birkhäuser Basel, 2000. DOI: `10.1007/978-3-0348-8424-2_1`.

[GMS06] J. L. Guermond, P. Minev and Jie Shen. An Overview of Projection Methods for Incompressible Flows. *Computer Methods in Applied Mechanics and Engineering*, 195(44-47):6011–6045, September 2006. DOI: `10.1016/j.cma.2005.10.010`.

[GR86] V. Girault and P.-A. Raviart. *Finite Element Methods for Navier-Stokes Equations.* Springer Berlin Heidelberg, 1986. DOI: `10.1007/978-3-642-61623-5`.

[He08] Y. He. The Euler Implicit/Explicit Scheme for the 2d Time-dependent Navier-Stokes Equations with Smooth or Non-smooth Initial Data. *Mathematics of Computation*, 77(264):2097–2124, May 2008. DOI: `10.1090/s0025-5718-08-02127-3`.

[HNW93] E. Hairer, S. P. Nørsett and G. Wanner. *Solving Ordinary Differential Equations I: Nonstiff Problems.* Springer Berlin Heidelberg, 2nd edition, 1993. DOI: `10.1007/978-3-540-78862-1`.

[HS07] Y. He and W. Sun. Stability and Convergence of the Crank-Nicolson/Adams-Bashforth Scheme for the Time-dependent Navier-Stokes Equations. *SIAM Journal on Numerical Analysis*, 45(2):837–869, January 2007. DOI: `10.1137/050639910`.

[HV03] W. Hundsdorfer and J. Verwer. *Numerical Solution of Time-dependent Advection-diffusion-reaction Equations.* Springer Berlin Heidelberg, 2003. DOI: `10.1007/978-3-662-09017-6`.

[HW08] J. S. Hesthaven and T. Warburton. *Nodal Discontinuous Galerkin Methods.* Springer New York, 2008. DOI: `10.1007/978-0-387-72067-8`.

[HW96] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II: Stiff and Differential-algebraic Problems.* Springer Berlin Heidelberg, 1st edition, 1996. DOI: `10.1007/978-3-642-05221-7`.

[JL04] H. Johnston and J.-G. Liu. Accurate, Stable and Efficient Navier-Stokes Solvers Based on Explicit Treatment of the Pressure Term. *Journal of Computational Physics*, 199(1):221–259, September 2004. DOI: `10.1016/j.jcp.2004.02.009`.

[JLM$^+$17] V. John, A. Linke, C. Merdon, M. Neilan and L. G. Rebholz. On the Divergence Constraint in Mixed Finite Element Methods for Incompressible Flows. *SIAM Review*, 59(3):492–544, 2017. DOI: `10.1137/15m1047696`.

[Joh16] V. John. *Finite Element Methods for Incompressible Flow Problems.* Springer Series in Computational Mathematics. Springer International Publishing, 4th November 2016. ISBN: 3319457497. DOI: `10.1007/978-3-319-45750-5`.

[JR10] V. John and J. Rang. Adaptive Time Step Control for the Incompressible Navier-Stokes Equations. *Computer Methods in Applied Mechanics and Engineering*, 199(9-12):514–524, January 2010. DOI: `10.1016/j.cma.2009.10.005`.

[Kan07] G. Kanschat. *Discontinuous Galerkin Methods for Viscous Incompressible Flow.* Advances in Numerical Mathematics. Teubner Research, 2007. ISBN: 978-3-8350-4001-4.

[KC03] C. A. Kennedy and M. H. Carpenter. Additive Runge-Kutta Schemes for Convection-diffusion-reaction Equations. *Applied Numerical Mathematics*, 44(1-2):139–181, January 2003. DOI: `10.1016/s0168-9274(02)00138-1`.

[KIO91] G. E. Karniadakis, M. Israeli and S. A. Orszag. High-order Splitting Methods for the Incompressible Navier-Stokes Equations. *Journal of Computational Physics*, 97(2):414–443, December 1991. DOI: `10.1016/0021-9991(91)90007-8`.

[KKW17] B. Krank, M. Kronbichler and W. A. Wall. Wall Modeling via Function Enrichment within a High-order DG Method for RANS Simulations of Incompressible Flow. *International Journal for Numerical Methods in Fluids*, 86(1):107–129, August 2017. DOI: `10.1002/fld.4409`.

[KM06] P. Kunkel and V. Mehrmann. *Differential-Algebraic Equations.* European Mathematical Society Publishing House, February 2006. DOI: `10.4171/017`.

[KM85] J. Kim and P. Moin. Application of a Fractional-step Method to Incompressible Navier-Stokes Equations. *Journal of Computational Physics*, 59(2):308–323, June 1985. DOI: `10.1016/0021-9991(85)90148-2`.

[KS05] G. Karniadakis and S. Sherwin. *Spectral/hp Element Methods for Computational Fluid Dynamics.* Oxford University Press, June 2005. DOI: `10.1093/acprof:oso/9780198528692.001.0001`.

[Lay08] W. Layton. *Introduction to the Numerical Analysis of Incompressible Viscous Flows.* Computational Science and Engineering. Society for Industrial and Applied Mathematic, 2008. ISBN: 978-0-898716-57-3.

[Leh10] C. Lehrenfeld. *Hybrid Discontinuous Galerkin Methods for Solving Incompressible Flow Problems.* Diplomarbeit, Rheinisch-Westfälischen Technischen Hochschule Aachen, May 2010. URL: `https://www.igpm.rwth-aachen.de/Download/reports/lehrenfeld/DA_HDG4NSE_1_0.pdf`.

[LRR00] S. Fabio Liotta, V. Romano and G. Russo. Central Schemes for Balance Laws of Relaxation Type. *SIAM Journal on Numerical Analysis*, 38(4):1337–1356, January 2000. DOI: `10.1137/s0036142999363061`.

[LS16] C. Lehrenfeld and J. Schöberl. High Order Exactly Divergence-free Hybrid Discontinuous Galerkin Methods for Unsteady Incompressible Flows. *Computer Methods in Applied Mechanics and Engineering*, 307:339–361, August 2016. DOI: `10.1016/j.cma.2016.04.025`.

[LS17] P. L. Lederer and J. Schöberl. Polynomial Robust Stability Analysis for H(div)-conforming Finite Elements for the Stokes Equations. *IMA Journal of Numerical Analysis*, August 2017. DOI: `10.1093/imanum/drx051`.

[MB01] A. J. Majda and A. L. Bertozzi. *Vorticity and Incompressible Flow.* Cambridge Texts in Applied Mathematics. Cambridge University Press, 2001. DOI: `10.1017/CBO9780511613203`.

[MPR90] Y. Maday, A. T. Patera and E. M. Rønquist. An Operator-integration-factor Splitting Method for Time-dependent Problems: Application to Incompressible Fluid Flow. *Journal of Scientific Computing*, 5(4):263–292, December 1990. DOI: `10.1007/bf01063118`.

[MT98] M. Marion and R. Temam. Navier-Stokes equations: Theory and approximation. In P. G. Ciarlet and J. L. Lions, editors, *Handbook of Numerical Analysis, Volume VI. Numerical Methods for Solids (Part 3) Numerical Methods for Fluids (Part 1)*, pages 503–689. Elsevier, 1998. DOI: `10.1016/s1570-8659(98)80010-0`.

[NL79] O. Nevanlinna and W. Liniger. Contractive methods for stiff differential equations part II. *BIT*, 19(1):53–72, March 1979. DOI: `10.1007/bf01931222`.

[PR00] L. Pareschi and G. Russo. *Implicit-Explicit Runge-Kutta Schemes for Stiff Systems of Differential Equations.* In *Recent Trends in Numerical Analysis.* D. Trigiante, editor. Volume 3. Nova Science Publishers, 2000, pages 269–289. ISBN: 1-56072-885-X.

[PR05] L. Pareschi and G. Russo. Implicit-Explicit Runge-Kutta Schemes and Applications to Hyperbolic Systems with Relaxation. *Journal of Scientific Computing*, 25(1):129–155, October 2005. DOI: `10.1007/s10915-004-4636-4`.

[Qin94] J. Qin. *On the Convergence of Some Low Order Mixed Finite Elements for Incompressible Fluids.* PhD thesis, Pennsylvania State University, 1994.

[Qua93] L. Quartapelle. *Numerical Solution of the Incompressible Navier-Stokes Equations.* Birkhäuser Basel, 1993. DOI: 10.1007/978-3-0348-8579-9.

[Riv08] B. Rivière. *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations: Theory and Implementation.* Frontiers in Applied Mathematics. Society for Industrial and Applied Mathematic, January 2008. DOI: 10.1137/1.9780898717440.

[Ruu93] S. J. Ruuth. *Implicit-Explicit Methods for Time-Dependent PDE's.* Master's Thesis, Institute of Applied Mathematics, University of British Columbia, 27th April 1993. DOI: 10.14288/1.0079818.

[Sch14] J. Schöberl. C++ 11 Implementation of Finite Elements in NGSolve, Preprint. *ASC Report 30/2014, Institute for Analysis and Scientific Computing, Vienna University of Technology*, 2014. URL: http://www.asc.tuwien.ac.at/~schoeberl/wiki/publications/ngs-cpp11.pdf.

[Sch97] J. Schöberl. NETGEN an Advancing Front 2d/3d-mesh Generator Based on Abstract Rules. *Computing and Visualization in Science*, 1(1):41–52, July 1997. DOI: 10.1007/s007910050004.

[Sch99] C. Schwab. *p- and hp- Finite Element Methods: Theory and Applications to Solid and Fluid Mechanics.* Numerical Mathematics and Scientific Computation. Clarendon Press, 1999. ISBN: 9780198503903.

[SJL^{+}18] P. W. Schroeder, V. John, P. L. Lederer, C. Lehrenfeld, G. Lube and J. Schöberl. On Reference Solutions and the Sensitivity of the 2d Kelvin-Helmholtz Instability Problem. 2018. arXiv: 1803.06893 [math.NA]. URL: http://arxiv.org/abs/1803.06893v2.

[SL17a] P. W. Schroeder and G. Lube. Divergence-free H(div)-FEM for Time-dependent Incompressible Flows with Applications to High Reynolds Number Vortex Dynamics. *Journal of Scientific Computing*, 75(2):830–858, September 2017. DOI: 10.1007/s10915-017-0561-1.

[SL17b] P. W. Schroeder and G. Lube. Pressure-robust Analysis of Divergence-free and Conforming FEM for Evolutionary Incompressible Navier-Stokes Flows. *Journal of Numerical Mathematics*, 25(4), December 2017. DOI: 10.1515/jnma-2016-1101.

[SLL^{+}18] P. W. Schroeder, C. Lehrenfeld, A. Linke and G. Lube. Towards Computable Flows and Robust Estimates for inf-sup Stable FEM Applied to the Time-dependent Incompressible Navier-Stokes Equations. *SeMA Journal*, April 2018. DOI: 10.1007/s40324-018-0157-1.

[SST02] D. Schötzau, C. Schwab and A. Toselli. Mixed hp-DGFEM for Incompressible Flows. *SIAM Journal on Numerical Analysis*, 40(6):2171–2194, January 2002. DOI: 10.1137/s0036142901399124.

[STD^{+}96] M. Schäfer, S. Turek, F. Durst, E. Krause and R. Rannacher. Benchmark Computations of Laminar Flow around a Cylinder. In *Notes on Numerical Fluid Mechanics (NNFM)*, pages 547–566. Vieweg+Teubner Verlag, 1996. DOI: 10.1007/978-3-322-89849-4_39.

[Tem77] R. Temam. *Navier-Stokes Equations: Theory and Numerical Analysis.* North-Holland, 1977. ISBN: 0-7204-2840-8.

[Ton04] F. Tone. Error Analysis for a Second Order Scheme for the Navier-Stokes Equations. *Applied Numerical Mathematics*, 50(1):93–119, July 2004. DOI: 10.1016/j.apnum.2003.12.003.

[WH03] T. Warburton and J. S. Hesthaven. On the Constants in hp-finite Element Trace Inverse Inequalities. *Computer Methods in Applied Mechanics and Engineering*, 192(25):2765–2773, June 2003. DOI: 10.1016/s0045-7825(03)00294-9.

[Wol18a] Wolfram|Alpha. 7th May 2018. URL: http://www.wolframalpha.com/input/?i=d%2Fdt+((e%5E(it)-3)(e%5E(it)-1)%2B(e%5E(-it)-3)(e%5E(-it)-1))%2F(2e%5E(-it)(2-e%5E(-it))).

[Wol18b] Wolfram|Alpha. 7th May 2018. URL: http://www.wolframalpha.com/input/?i=d%2Fdt+(e%5E(it)(2-e%5E(it)))%2F(e%5E(-it)(2-e%5E(-it))).

[Wol18c] Wolfram|Alpha. 8th May 2018. URL: http://www.wolframalpha.com/input/?i=d%2Fdt+((16(1-e%5E(it)))%2F(e%5E(2it)%2B6e%5E(it)%2B9)%2B(16(1-e%5E(-it)))%2F(e%5E(-2it)%2B6e%5E(-it)%2B9))%2F((8e%5E(-it)(3-e%5E(-it)))%2F(e%5E(-2it)%2B6e%5E(-it)%2B9)).

[Wol18d] Wolfram|Alpha. 8th May 2018. URL: http://www.wolframalpha.com/input/?i=d%2Fdt+((8e%5E(it)(3-e%5E(it)))%2F(e%5E(2it)%2B6e%5E(it)%2B9))%2F((8e%5E(-it)(3-e%5E(-it)))%2F(e%5E(-2it)%2B6e%5E(-it)%2B9)).

[Wol18e] Wolfram|Alpha. 8th May 2018. URL: http://www.wolframalpha.com/input/?i=d%2Fdt+(((4*(1-e%5E(it)))%2F(e%5E(2it)%2B3))%2B((4(1-e%5E(-it)))%2F(e%5E(-2it)%2B3)))%2F((2e%5E(-it)(3-e%5E(-it)))%2F(e%5E(-2it)%2B3)).

[Wol18f] Wolfram|Alpha. 8th May 2018. URL: http://www.wolframalpha.com/input/?i=d%2Fdt+((2e%5E(it)(3-e%5E(it)))%2F(e%5E(2it)%2B3))%2F((2e%5E(-it)(3-e%5E(-it)))%2F(e%5E(-2it)%2B3)).

[Zha04] S. Zhang. A New Family of Stable Mixed Finite Elements for the 3d Stokes Equations. *Mathematics of Computation*, 74(250):543–555, August 2004. DOI: 10.1090/s0025-5718-04-01711-9.

Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel verfasst habe.

Göttingen, den 7. August 2018 UNTERSCHRIFT