Karlsruhe Institute for Technology

Institute for Applied and Numerical Mathematics

Master Thesis

Error Analysis for Full Discretizations of Maxwell's Equations with Explicit Runge-Kutta Methods

Andreas Sturm

Supervisor: Prof. Dr. Marlis Hochbruck

Acknowledgement

I would like to sincerely thank my supervisor, Prof. Dr. Marlis Hochbruck, for giving me the chance to work as a HiWi in the Research Training Group 1294: ”Analysis, Simulation and Design of Nanotechnological Processes” of the German Research Foundation (DFG). This work built the foundation for my master’s thesis and I am grateful to Prof. Hochbruck for constant and great support on it. She spent plenty of time discussing my problems and ideas and thus made this thesis possible.

Also, I want to ensure my thanks to Dr. Tomislav Pažur who dedicated many hours to introducing me into his research project on dG methods and implicit time integrators. Many ideas in this thesis are motivated by his supervising as well as the numerical results are based on his work on matlab codes for discretizing Maxwell’s equations with dG methods. Furthermore, I am very grateful to Tomislav for his proofreading of my thesis and his numerous advises on it.

Contents

Introduction

This thesis is concerned with the numerical analysis of time-dependent linear Maxwell's equations. We follow the method of lines ansatz where we first discretize Maxwell's equations in space yielding a (large) system of ODEs which are subsequently solved using a time integration method.

For the spatial discretization we use discontinuous Galerkin (dG) finite element methods which has become of great interest in recent years, see the textbooks [8, 17]. The popularity of dG methods relies in several advantages compared to finite differences (FD) or standard finite element (FE) methods. The main benefits with respect to FD methods are that dG methods can handle irregular domains and admit high-order accuracy as well as adaptivity. In view of FE methods the superior aspects of dG methods are that they can handle more easily non-conforming meshes, the mass matrix is block-diagonal and they are accessible for high parallelization. Furthermore, dG methods are well-suited for solving Maxwell's equations in composite media, i. e. media with piecewise constant material coefficients.

The dG discretization of Maxwell's equations results in a semi-discrete problem which corresponds to a system of ODEs and which yet has to be integrated in time. We therefore work with explicit Runge-Kutta (RK) methods with one, two or three stages. Convergence results for dG methods combined with two- and three-stage RK methods have been proven in 2010 by Burman, Ern and Fernández, see [2]. The time integration analysis in this thesis strongly relies on this paper but we propose a modified notation and generalize some aspects. It is characteristic for explicit time-integration methods that the step size has to be restricted due to stability requirements (CFL condition). One can overcome this problem by considering implicit or exponential time integrators and we refer the reader to [16] for this methods. However, this methods require to solve large systems of linear equations or to evaluate matrix exponentials for large matrices, which complicates the implementation. In contrary, explicit time integrators can exploit the block-diagonal structure of the mass-matrix and thereby lead to fully explicit schemes.

Outline of the Thesis

We organize this thesis as follows. In Chapter 1 we introduce Maxwell's equation and give the physical interpration of the appearing quantities. Then, we turn to linear Maxwell's equations in an inhomogeneous, isotropic medium. We conclude the chapter by providing a mathematical framework to proof well-posedness of Maxwell's equations.

Subsequently, we discretize Maxwell's equations in space. Therefore, we introduce in Chapter 2 the discrete setting dG methods are based upon. In Chapter 3 we derive the dG discretization of Maxwell's equations. The stability analysis of the dG discretization will enable us to prove convergence of order k, where k denotes the polynomial degree used in the dG method. A further analysis of a so called stabilized dG discretization will then allow us to improve the convergence order to $k + 1/2$.

Chapter 4 is dedicated to full discretizations of Maxwell's equations stemming from discretizing the semi-discrete problem provided by the dG spatial discretization with explicit RK methods. Thereby, we start by introducing (explicit) RK methods. Our further analysis is based on energy techniques and we first deduce energy identities associated with the RK approximations. This will allow us to prove the stability of the full discretization and then lead us to proof convergence of order k in space and s in

time, when s denotes the number of stages of the RK scheme. Finally, we prove convergence of order $k + 1/2$ in space and s in time for stabilized dG methods.

In the end, Chapter 5 provides some implementational aspects of dG discretizations and then concludes the thesis by illustrating the gained results by numerical experiments.

Chapter 1

Maxwell's Equations

In this chapter we give an introduction into Maxwell's equations where we mainly follow [5, 12, 15, 16]. At first we state Maxwell's equations in their general differential form. Subsequently, we consider the particular case of linear Maxwell's equations in an inhomogeneous, isotropic material which is surrounded by a perfect conductor.

Thereafter, we introduce a mathematical framework in which Maxwell's equations can be embedded. We show that in this context Maxwell's equations can be stated as an abstract evolution equation. Finally, we ensure the well-posedness of this evolution equation with the theory of C_0-groups, in particular with Stone's theorem. This guarantees the existence of a unique solution of the linear Maxwell's equations.

1.1 The Partial Differential Equations

We introduce the vector fields $\mathcal{D}, \mathcal{E}, \mathcal{B}, \mathcal{H} : \mathbb{R}_+ \times \Omega \to \mathbb{R}^3$ on a set $\Omega \subset \mathbb{R}^3$, where $\mathcal{D}$ represents the *electric displacement field*, $\mathcal{E}$ the *electric field*, $\mathcal{H}$ the *magnetic induction* and $\mathcal{B}$ the *magnectic field intensity*. Then, *Maxwell's equations* can be stated as

$$\partial_t \mathcal{B} + \nabla \times \mathcal{E} = 0, \tag{1.1a}$$
$$\partial_t \mathcal{D} - \nabla \times \mathcal{H} = -\mathcal{J}, \tag{1.1b}$$
$$\nabla \cdot \mathcal{D} = \rho, \tag{1.1c}$$
$$\nabla \cdot \mathcal{B} = 0, \tag{1.1d}$$

for given *electric current density* $\mathcal{J} : \mathbb{R}_+ \times \Omega \to \mathbb{R}^3$ and *electric charge density* $\rho : \mathbb{R}_+ \times \Omega \to \mathbb{R}$.

The first equation is called *Faraday's law of induction* and describes the effect of a time-varying magnetic field on the electric field. The second equation is *Ampère's law* and states the effect of the (external and internal) current on the magnetic field. The last two equations are *Gauss's electric law* and *Gauss's magnetic law*, respectively. The former describes the sources of the electric displacement whereas the latter states that there are no magnetic currents. For a deeper insight in the physics of Maxwell's equations we refer to [11].

A result that follows immediately from above equations is *conservation of charge*, i. e., there holds

$$\partial_t \rho + \nabla \cdot \mathcal{J} = 0. \tag{1.2}$$

Indeed, we see this by differentiating (1.1c) with respect to (w. r. t.) t and plugging it into (1.1b). The result then follows by the identity $\nabla \cdot (\nabla \times \cdot) = 0$.

An additional result concerns the time evolution of Maxwell's equations. Let us therefore consider the last two equations, often called the *div-equations* owing to the derivatives they contain. Differenti-

ating w. r. t. t yields

$$\partial_t(\nabla \cdot \mathcal{D} - \rho) = \nabla \cdot (\nabla \times \mathcal{H} - \mathcal{J}) - \partial_t \rho = 0,$$
$$\partial_t \nabla \cdot \mathcal{B} = -\nabla \cdot (\nabla \times \mathcal{E}) = 0,$$

where we have used the first two equations (the *curl-equations*) (1.1a)-(1.1b), the conservation of charge (1.2) and the identity $\nabla \cdot (\nabla \times \cdot) = 0$. Thus, if the div-equations are satisfied for some initial time they will be fulfilled for every time. This essentially means that in order to analyze the time evolution of Maxwell's equations it is sufficient to consider the two curl-equations (1.1a)-(1.1b).

1.2 The Constitutive Equations

As we have seen, Maxwell's equations consist of six independet scalar equations for twelve scalar unknowns and are thus not consistent. This is overcome by introducing the so-called *constitutive equations* which couple the fields by

$$\mathcal{D} = \mathcal{D}(\mathcal{E}, \mathcal{H}), \qquad \mathcal{B} = \mathcal{B}(\mathcal{E}, \mathcal{H}).$$

For example, we have in vacuum

$$\mathcal{D} = \varepsilon_0 \mathcal{E}, \qquad \mathcal{B} = \mu_0 \mathcal{H},$$

where ε_0 is the *permittivity of free space* and μ_0 the *permeability of free space*. These constants are related to the *speed of light* in vacuum, here called c_0, by

$$c_0 = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}.$$

In material the situation can be more complicated. For example in *inhomogeneous* and *anisotropic* media we can model the dependencies of the fields by *linear constitutive equations* of the form

$$\mathcal{D} = \varepsilon \mathcal{E}, \qquad \mathcal{B} = \mu \mathcal{H},$$

with matrix-valued functions $\varepsilon : \mathbb{R}^3 \to \mathbb{R}^{3\times 3}$ and $\mu : \mathbb{R}^3 \to \mathbb{R}^{3\times 3}$ called the *permittivity tensor* and *permeability tensor*, respectively. For the rest of the thesis we deal with *isotropic* media, where the permittivity and the permeability are directionally independent. In this case the constitutive equations simplify to

$$\mathcal{D} = \varepsilon_0 \varepsilon_r \mathcal{E}, \qquad \mathcal{B} = \mu_0 \mu_r \mathcal{H}, \tag{1.3}$$

where ε_r, $\mu_r : \mathbb{R}^3 \to \mathbb{R}_+$ are scalar functions called the *relative permittivity* and the *relative permeability* of the medium. We set $\varepsilon = \varepsilon_0 \varepsilon_r$ and $\mu = \mu_0 \mu_r$ and refer to them as the *permittivity* and the *permeability* of the medium.

Furthermore, the current density $\mathcal{J}$ can depend on the material and on the fields. For conducting media this can be modeled by *Ohm's law*:

$$\mathcal{J} = \sigma \mathcal{E} + \mathcal{J}_e. \tag{1.4}$$

Here $\sigma : \mathbb{R}^3 \to \mathbb{R}$ is the *conductivity* and $\mathcal{J}_e : \mathbb{R}^3 \to \mathbb{R}^3$ is the *external current density*.

1.3 Linear Maxwell's Equations

So far, we have introduced Maxwell's equations in an inhomogeneous, isotropic material. In order to study the time evoluton of Maxwell's equations in a bounded domain Ω, we have to introduce suitable boundary conditions on $\partial\Omega$.

Boundary conditions We consider the case where the material is surrounded by a perfect conductor. In [5] it is shown that this yields the boundary conditions

$$n \times \mathcal{E} = 0, \qquad n \cdot (\mu \mathcal{H}) = 0, \tag{1.5}$$

where n denotes the outward unit normal to $\partial\Omega$.

Linear Maxwell's equations Incorporating the constitutive equations and the boundary conditions into Maxwell's equations results in the following *evolution problem*: Given the current density $\mathcal{J}$, the charge density ρ, and the initial values $\mathcal{E}_0$, $\mathcal{H}_0$, we search for the electric and magnetic field $\mathcal{E}, \mathcal{H} : \mathbb{R}_+ \times \Omega \to \mathbb{R}^3$, such that

$$\begin{aligned}
\mu \partial_t \mathcal{H} + \nabla \times \mathcal{E} &= 0 && \text{in } \mathbb{R}_+ \times \Omega, && \text{(1.6a)}\\
\varepsilon \partial_t \mathcal{E} - \nabla \times \mathcal{H} &= -\mathcal{J} && \text{in } \mathbb{R}_+ \times \Omega, && \text{(1.6b)}\\
\nabla \cdot (\varepsilon \mathcal{E}) &= \rho, && \text{in } \mathbb{R}_+ \times \Omega, && \text{(1.6c)}\\
\nabla \cdot (\mu \mathcal{H}) &= 0, && \text{in } \mathbb{R}_+ \times \Omega, && \text{(1.6d)}\\
n \times \mathcal{E} &= 0 && \text{on } \mathbb{R}_+ \times \partial\Omega, && \text{(1.6e)}\\
n \cdot (\mu \mathcal{H}) &= 0 && \text{on } \mathbb{R}_+ \times \partial\Omega, && \text{(1.6f)}\\
\mathcal{E}(t=0) &= \mathcal{E}_0 && \text{in } \Omega, && \text{(1.6g)}\\
\mathcal{H}(t=0) &= \mathcal{H}_0 && \text{in } \Omega. && \text{(1.6h)}
\end{aligned}$$

Reduced linear Maxwell's equations We already carried out that the div-equations (1.6c)-(1.6d) can be dropped when analyzing the time-evolution. Later we will see that the same holds true for the second boundary condition (1.6f), i. e. it is automatically satisfied if it is satisfied for some initial time. Thus, it is suffcient to consider following reduced problem: We search for $\mathcal{E}, \mathcal{H} : \mathbb{R}_+ \times \Omega \to \mathbb{R}^3$, such that

$$\begin{aligned}
\mu \partial_t \mathcal{H} + \nabla \times \mathcal{E} &= 0 && \text{in } \mathbb{R}_+ \times \Omega, && \text{(1.7a)}\\
\varepsilon \partial_t \mathcal{E} - \nabla \times \mathcal{H} &= -\mathcal{J} && \text{in } \mathbb{R}_+ \times \Omega, && \text{(1.7b)}\\
n \times \mathcal{E} &= 0 && \text{on } \mathbb{R}_+ \times \partial\Omega, && \text{(1.7c)}\\
\mathcal{E}(t=0) &= \mathcal{E}_0 && \text{in } \Omega, && \text{(1.7d)}\\
\mathcal{H}(t=0) &= \mathcal{H}_0 && \text{in } \Omega. && \text{(1.7e)}
\end{aligned}$$

We conclude this section with a special case of Maxwell's equation which admits to reduce the three dimensional system (1.7) to a two dimensional problem. This is of particular interest, since our later numerical experiments are carried out for this case.

1.3.1 Reduction to Two Dimensions

In [15] it is pointed out that if the underlying physical system is homogeneous in z-direction Maxwell's equations (1.7) decouple into two sets of three equations.

TE polarization The first set reads

$$\begin{aligned}
\varepsilon \partial_t \mathcal{E}_x - \partial_y \mathcal{H}_z &= -\mathcal{J}_x && \text{in } \mathbb{R}_+ \times \Omega, && \text{(1.8a)}\\
\varepsilon \partial_t \mathcal{E}_y + \partial_x \mathcal{H}_z &= -\mathcal{J}_y && \text{in } \mathbb{R}_+ \times \Omega, && \text{(1.8b)}\\
\mu \partial_t \mathcal{H}_z + \partial_x \mathcal{E}_y - \partial_y \mathcal{E}_x &= 0 && \text{in } \mathbb{R}_+ \times \Omega, && \text{(1.8c)}\\
n_x \mathcal{E}_y - n_y \mathcal{E}_x &= 0 && \text{in } \mathbb{R}_+ \times \partial\Omega, && \text{(1.8d)}
\end{aligned}$$

and describes the so-called *transverse-electric (TE) polarization*. In this case the electric field lies in the plane of propagation.

TM polarization In contrary, the second set, called *transverse-magnetic (TM) polarization*,

$$\mu\partial_t\mathcal{H}_x + \partial_y\mathcal{E}_z = 0 \qquad \text{in } \mathbb{R}_+ \times \Omega, \tag{1.9a}$$
$$\mu\partial_t\mathcal{H}_y - \partial_x\mathcal{E}_z = 0 \qquad \text{in } \mathbb{R}_+ \times \Omega, \tag{1.9b}$$
$$\varepsilon\partial_t\mathcal{E}_z + \partial_y\mathcal{H}_x - \partial_x\mathcal{H}_y = -\mathcal{J}_z \qquad \text{in } \mathbb{R}_+ \times \Omega, \tag{1.9c}$$
$$\mathcal{E}_z = 0 \qquad \text{in } \mathbb{R}_+ \times \partial\Omega, \tag{1.9d}$$

desbribes an electric field perpendicular to the plane of of propagation.

1.4 Mathematical Aspects of Maxwell's Equations

The later discretization is tailored to approximate Maxwell's equations in an L^2-sense and thus the aim of this section is to formulate Maxwell's equations (1.7) as a well-posed mathematical problem in an L^2 setting. This gives rise to the question of the meaning of the curl-operator $\nabla\times$ in (1.7a)-(1.7b) since L^2-functions do not necessarily posses (weak) derivatives. Furthermore, we have to give a meaning to the boundary condition (1.7c). Thus, we will at first introduce the concept of the so called *graph space*, where the curl of a function is defined in a variational way. A further restriction of the graph space will then allow us to incorporate the boundary condition.

Throughout this thesis we consider only bounded domains Ω which posses a Lipschitz-continuous boundary $\partial\Omega$. Then, the *outward unit normal* n is defined almost everywhere (a. e.) on Ω. Furthermore, we suppose that the coefficients ε and μ, are in $L^\infty(\Omega)$ and that they are uniformly positive, meaning that there is a constant $\delta > 0$ such that

$$\varepsilon, \mu \geq \delta > 0. \tag{1.10}$$

1.4.1 The State Space

We begin by introducing the basic space.

Definition 1.1 (The state space V). We define the *state space* V as

$$V := L^2(\Omega)^3 \times L^2(\Omega)^3. \tag{1.11}$$

We equip it with the inner product: For $[H_1, E_1]^T, [H_2, E_2]^T \in V$,

$$\left(\begin{bmatrix} H_1 \\ E_1 \end{bmatrix}, \begin{bmatrix} H_2 \\ E_2 \end{bmatrix}\right)_V := \int_\Omega \mu H_1 \cdot H_2 + \varepsilon E_1 \cdot E_2, \tag{1.12}$$

and with the associated norm: For $[H, E]^T \in V$,

$$\left\|\begin{bmatrix} H \\ E \end{bmatrix}\right\|_V = \left(\begin{bmatrix} H \\ E \end{bmatrix}, \begin{bmatrix} H \\ E \end{bmatrix}\right)_V^{1/2}. \tag{1.13}$$

Owing to the positivity assumption (1.10) the V-inner product is equivalent to the standard L^2-inner product and thus $(V, (\cdot,\cdot)_V)$ is a Hilbert space. The usage of the V-inner product instead of the standard L^2-inner product is motivated by its physical meaning. In fact, its induced norm represents the *electromagnetic energy*.

1.4.2 The Graph Space

Now let us give a meaning to the curl operator $\nabla\times$ in (1.7a)-(1.7b). We cannot use the standard curl operator since it is only defined for continuous differentiable functions and clearly functions in V do not have to satisfy this. Recalling that we are working in an L^2-setting we see that for a function $[H,E]^T \in V$ the formal condition $\nabla\times H$, $\nabla\times E \in L^2(\Omega)^3$ is sufficient to guarantee that (1.7a)-(1.7b) are well-defined. So let us define what the statement $\nabla\times F \in L^2(\Omega)^3$ means here. Let therefore $C_0^\infty(\Omega)^3$ denote the space of infinitely differentiable functions with compact support in Ω.

Definition 1.2 (The variational curl). We say that a function $F \in L^2(\Omega)^3$ posses a *variational curl* in $L^2(\Omega)^3$ if there exists a function $G \in L^2(\Omega)^3$ such that

$$\int_\Omega G\cdot\varphi = \int_\Omega F\cdot(\nabla\times\varphi) \qquad \forall\varphi\in C_0^\infty(\Omega)^3. \tag{1.14}$$

We write $\nabla\times F = G$.

Since the space $C_0^\infty(\Omega)^3$ is dense in $L^2(\Omega)^3$ [19, Lemma V.1.10] the variational curl is unique (if it exists).

Remark 1.3 This definition is a generalization of an integration by parts formula. Indeed, for a weakly differentiable function $\widetilde{F}\in H^1(\Omega)^3$, there holds

$$\int_\Omega(\nabla\times\widetilde{F})\cdot\varphi = \int_\Omega \widetilde{F}\cdot(\nabla\times\varphi), \qquad \forall\varphi\in C_0^\infty(\Omega)^3.$$

◇

We collect functions admitting a variational curl in the following space.

Definition 1.4 (The graph space $H(\mathrm{curl},\Omega)$). The *graph space* of the curl-operator is defined as

$$H(\mathrm{curl},\Omega) := \{F\in L^2(\Omega)^3 \,|\, \nabla\times F\in L^2(\Omega)^3\}, \tag{1.15}$$

with the natural inner product: For F, $G\in H(\mathrm{curl},\Omega)$,

$$(F,G)_{H(\mathrm{curl},\Omega)} := (F,G)_{L^2(\Omega)^3} + (\nabla\times F,\nabla\times G)_{L^2(\Omega)^3}, \tag{1.16}$$

and the associoated *graph norm* $\|F\|_{H(\mathrm{curl},\Omega)} := (F,F)^{1/2}_{H(\mathrm{curl},\Omega)}$.

Proposition 1.5 *The graph space $H(\mathrm{curl},\Omega)$ is a Hilbert space.*

Proof: Let (F_n) be a Cauchy sequence in $H(\mathrm{curl},\Omega)$. Then, (F_n) and $(\nabla\times F_n)$ are Cauchy sequences in $L^2(\Omega)^3$ and thus convergent. Let F and G denote their respective limits in $L^2(\Omega)^3$. Then, for all $\varphi\in C_0^\infty(\Omega)$ and all $n\in\mathbb{N}$ we have by the definition of $H(\mathrm{curl},\Omega)$,

$$\int_\Omega(\nabla\times F_n)\cdot\varphi = \int_\Omega F_n\cdot(\nabla\times\varphi).$$

Taking the limit $n\to\infty$ yields

$$\int_\Omega G\cdot\varphi = \int_\Omega F\cdot(\nabla\times\varphi).$$

Thus, we conclude from (1.14) that $F\in H(\mathrm{curl},\Omega)$ and $G=\nabla\times F$. □

The following result will be important.

Theorem 1.6 *([14, Theorem 3.26]). The space $H(\mathrm{curl},\Omega)$ is the closure of $C^\infty(\overline{\Omega})^3$ w. r. t. the graph norm $\|\cdot\|_{H(\mathrm{curl},\Omega)}$.*

1.4.3 Boundary Conditions in the Graph Space

Next, we incorporate the boundary condition (1.7c) in the graph space $H(\mathrm{curl},\Omega)$. Motivated by Theorem 1.6 we establisch the desired property as follows:

Definition 1.7 (The space $H_0(\mathrm{curl},\Omega)$). We define the space $H_0(\mathrm{curl},\Omega)$ as the closure of $C_0^\infty(\Omega)^3$ w. r. t. the graph norm $\|\cdot\|_{H(\mathrm{curl},\Omega)}$.

We easily see that $H_0(\mathrm{curl},\Omega)$ is a closed subspace of the Hilbert space $H(\mathrm{curl},\Omega)$ and thus a Hilbert space itself. The next two lemmas clearify the properties of functions in $H_0(\mathrm{curl},\Omega)$.

Lemma 1.8 *([12, Lemma 4.17]). The space $\{\nabla v \,|\, v \in H_0^1(\Omega)\}$ is a closed subspace of $H_0(\mathrm{curl},\Omega)$.*

This lemma essentially states that functions belonging to $H_0(\mathrm{curl},\Omega)$ do not need to have a vanishing normal component on the boundary $\partial\Omega$. For example take $\Omega = [-1,1]^2$ and $v = (x^2-1)(y^2-1)$. Then, clearly it holds $v \in H_0^1(\Omega)$ and Lemma 1.8 applies yielding $\nabla v \in H_0(\mathrm{curl},\Omega)$. However, we have on the boundary

$$\nabla v \cdot n = \begin{cases} \pm 2x(y^2-1), & \text{for } (x,y) \in \{\pm 1\} \times [-1,1], \\ \pm 2y(x^2-1), & \text{for } (x,y) \in [-1,1] \times \{\pm 1\}, \end{cases}$$

which obviously does not vanish. In contrary, observe that the tangential component of ∇v vanishes since

$$\nabla v \times n = \begin{cases} \pm 2y(x^2-1), & \text{for } (x,y) \in \{\pm 1\} \times [-1,1], \\ \pm 2x(y^2-1), & \text{for } (x,y) \in [-1,1] \times \{\pm 1\}. \end{cases}$$

Indeed, this holds true for all function belonging to $H_0(\mathrm{curl},\Omega)$, namely, their tangential component has to vanish. The proof can be found in [14, Theorem 3.29] and requires the introduction of the space $H^{-1/2}(\partial\Omega)^3$, which we ommit here. We rather cite following two lemmas illustrating this result.

Lemma 1.9 *([12, Lemma 4.18]). The space $\{F \in C^1(\overline{\Omega})^3 \,|\, n \times F = 0 \text{ on } \partial\Omega\}$ is a subspace of $H_0(\mathrm{curl},\Omega)$.*

Lemma 1.10 *([14, Lemma 3.27]). Let $F \in H(\mathrm{curl},\Omega)$ be such that for every $\varphi \in C^\infty(\overline{\Omega})^3$ it holds*

$$(\nabla \times F, \varphi)_{L^2(\Omega)^3} = (F, \nabla \times \varphi)_{L^2(\Omega)^3}. \tag{1.17}$$

Then, $F \in H_0(\mathrm{curl},\Omega)$.

Remark 1.11 Let us point out the statement of Lemma 1.10 for functions with more regularity, say $\widetilde{F} \in H^1(\Omega)^3$. Integration by parts gives

$$\int_\Omega (\nabla \times \widetilde{F}) \cdot \varphi = \int_\Omega \widetilde{F} \cdot (\nabla \times \varphi) + \int_{\partial\Omega} (n \times \widetilde{F}) \cdot \varphi, \qquad \forall \varphi \in C^\infty(\overline{\Omega})^3.$$

Owing to (1.17), we have

$$\int_{\partial\Omega} (n \times \widetilde{F}) \cdot \varphi = 0 \quad \forall \varphi \in C^\infty(\overline{\Omega})^3.$$

Since $C^\infty(\overline{\Omega})^3$ is dense in $L^2(\Omega)^3$, see [9], this is equivalent to

$$n \times \widetilde{F} = 0 \text{ a. e. on } \partial\Omega.$$

◇

We conclude this section with a generalization of *Green's theorem* to functions in $H(\mathrm{curl},\Omega)$.

Lemma 1.12 *(Green's theorem, [14, Lemma 3.27]). Let $H \in H(\mathrm{curl},\Omega)$ and $E \in H_0(\mathrm{curl},\Omega)$. Then, it holds*

$$(H, \nabla \times E)_{L^2(\Omega)^3} = (\nabla \times H, E)_{L^2(\Omega)^3}. \tag{1.18}$$

1.4.4 Well-Posedness

The results from the previous section allow us to collect the curl terms in (1.7a), (1.7b) as well as the associated boundary condition (1.7c) in an operator.

Definition 1.13 (Maxwell operator). We define the *Maxwell operator* A as

$$A : \mathcal{D}(A) \to V, \quad \begin{bmatrix} H \\ E \end{bmatrix} \mapsto \begin{bmatrix} \mu^{-1}\nabla \times E \\ -\varepsilon^{-1}\nabla \times H \end{bmatrix}, \tag{1.19}$$

where the *domain* of A is given as

$$\mathcal{D}(A) := H(\mathrm{curl}, \Omega) \times H_0(\mathrm{curl}, \Omega). \tag{1.20}$$

We endow $\mathcal{D}(A)$ with the graph norm: For $[H, E]^T \in \mathcal{D}(A)$,

$$\left\| \begin{bmatrix} H \\ E \end{bmatrix} \right\|_A^2 := \left\| \begin{bmatrix} H \\ E \end{bmatrix} \right\|_V^2 + \left\| A \begin{bmatrix} H \\ E \end{bmatrix} \right\|_V^2. \tag{1.21}$$

Clearly, there holds

$$\left\| \begin{bmatrix} H \\ E \end{bmatrix} \right\|_A^2 = \left\| \begin{bmatrix} \mu^{1/2} H \\ \varepsilon^{1/2} E \end{bmatrix} \right\|_{L^2(\Omega)^6}^2 + \left\| \begin{bmatrix} \mu^{-1/2}\nabla \times E \\ \varepsilon^{-1/2}\nabla \times H \end{bmatrix} \right\|_{L^2(\Omega)^6}^2.$$

Recalling that the coefficients μ and ε are assumend to be bounded and to be uniformly positive, see (1.10), we deduce the equivalence of the norms $\|\cdot\|_A$ and $\|\cdot\|_{H(\mathrm{curl},\Omega)\times H(\mathrm{curl},\Omega)}$. Furthermore, we can conclude that $(\mathcal{D}(A), \|\cdot\|_A)$ is a Hilbert space and consequently that A is a closed operator.

Homogeneous evolution equation We first consider the *homogeneous* case of Maxwell's equations (1.7), where the electric current and the electric charge density are zero, $\mathcal{J} = \rho = 0$. Then, we can rewrite(1.7) in a more compact form, namely as the abstract *evolution equation*: For a given initial value $u_0 = [\mathcal{H}_0, \mathcal{E}_0]^T \in \mathcal{D}(A)$ we search for $u = [\mathcal{H}, \mathcal{E}]^T \in C^1(\mathbb{R}_+; V) \cap C(\mathbb{R}_+; \mathcal{D}(A))$ such that

$$\partial_t u + Au = 0, \qquad t \geq 0, \tag{1.22a}$$

$$u(0) = u_0. \tag{1.22b}$$

Our aim is to show well-posedness of (1.22) via Stone's theorem A.1. Therefore, we show that its premises are satisfied, i. e. the domain $\mathcal{D}(A)$ is dense in V and the operator A is skew-adjoint. For the first assumption we notice that $C^\infty(\overline{\Omega})^3 \times C_0^\infty(\Omega)^3$ is a subset of $\mathcal{D}(A)$ (see Theorem 1.6 and Definition 1.7). Then, the assumption readily follows by the density of both $C^\infty(\Omega)^3$ and $C_0^\infty(\Omega)^3$ in $L^2(\Omega)^3$ w. r. t. to the L^2-norm and the equivalence of the V-norm and the L^2-norm. Thus, it remains to show skew-adjointness of A.

Proposition 1.14 *(Skew-adjointness of A). The Maxwell operator A is skew-adjoint w. r. t. to the V-inner product.*

Proof: We follow the proof in [16, Proposition 2.1]. We have to show that the domain of A and the domain of its adjoint $A^\star$ coincide, $\mathcal{D}(A) = \mathcal{D}(A^\star)$, and that A is skew-symmetric, i. e. for all $v_1, v_2 \in \mathcal{D}(A)$ it holds

$$(Av_1, v_2)_V = -(v_1, Av_2)_V.$$

Let us begin by proving that A is skew-symmetric. For $v_1 = [H_1, E_1]^T, v_2 = [H_2, E_2]^T \in \mathcal{D}(A)$ there holds

$$\begin{aligned}(Av_1, v_2)_V &= \left(\begin{bmatrix} \mu^{-1}\nabla \times E_1 \\ -\varepsilon^{-1}\nabla \times H_1 \end{bmatrix}, \begin{bmatrix} H_2 \\ E_2 \end{bmatrix}\right)_V \\ &= \left(\begin{bmatrix} \nabla \times E_1 \\ -\nabla \times H_1 \end{bmatrix}, \begin{bmatrix} H_2 \\ E_2 \end{bmatrix}\right)_{L^2(\Omega)^6} \\ &= (\nabla \times E_1, H_2)_{L^2(\Omega)^3} - (\nabla \times H_1, E_2)_{L^2(\Omega)^3} \\ &= (E_1, \nabla \times H_2)_{L^2(\Omega)^3} - (H_1, \nabla \times E_2)_{L^2(\Omega)^3} \\ &= -\left(\begin{bmatrix} H_1 \\ E_1 \end{bmatrix}, \begin{bmatrix} \mu^{-1}\nabla \times E_2 \\ -\varepsilon^{-1}\nabla \times H_2 \end{bmatrix}\right)_V = -(v_1, Av_2)_V.\end{aligned}$$

Here we used Green's theorem (1.12) in the fourth line.
We continue by proving the coincidence of the domains of A and $A^\star$. The domain of the adjoint $A^\star$ is given by

$$\mathcal{D}(A^\star) = \{v_2 \in V \,|\, \exists v_3 \in V \text{ s.t. } \forall v_1 \in \mathcal{D}(A) : (Av_1, v_2)_V = (v_1, v_3)_V\}.$$

We show that both domains contain each other, $\mathcal{D}(A) \subset \mathcal{D}(A^\star)$ and $\mathcal{D}(A^\star) \subset \mathcal{D}(A)$. The first inclusion immediately follows with the computations above: Let $v_2 \in \mathcal{D}(A)$ and set $v_3 = -Av_2$. Then, for all $v_1 \in \mathcal{D}(A)$ there holds

$$(Av_1, v_2)_V = (v_1, v_3),$$

and thus $\mathcal{D}(A) \subset \mathcal{D}(A^\star)$. Conversely let $v_2 = [H_2, E_2]^T \in \mathcal{D}(A^\star)$. Then, by the definition of $\mathcal{D}(A^\star)$, there is $v_3 = [H_3, E_3]^T \in V$ such that for all $v_1 = [H_1, E_1]^T \in \mathcal{D}(A)$ it holds,

$$(Av_1, v_2)_V = (v_1, v_3)_V,$$

or equivalently,

$$(\nabla \times E_1, H_2)_{L^2(\Omega)^3} - (\nabla \times H_1, E_2)_{L^2(\Omega)^3} = (\mu H_1, H_3)_{L^2(\Omega)^3} + (\varepsilon E_1, E_3)_{L^2(\Omega)^3}. \tag{1.23}$$

We choose $H_1 = 0$,

$$(\nabla \times E_1, H_2)_{L^2(\Omega)^3} = (\varepsilon E_1, E_3)_{L^2(\Omega)^3}.$$

Since this holds for all $E_1 \in H_0(\text{curl}, \Omega)$ it holds also for all $E_1 \in C_0^\infty(\Omega)^3$. Thus, we have

$$\int_\Omega (\nabla \times E_1) \cdot H_2 = \int_\Omega \varepsilon E_1 \cdot E_3 \quad \forall E_1 \in C_0^\infty(\Omega)^3.$$

Recalling the definition of the variational curl (1.14) we conclude that $\nabla \times H_2 = \varepsilon E_3 \in L^2(\Omega)^3$ and therefore $H_2 \in H(\text{curl}, \Omega)$.
Similar, by choosing $E_1 = 0$ in (1.23), we get

$$-(\nabla \times H_1, E_2)_{L^2(\Omega)^3} = (\mu H_1, H_3)_{L^2(\Omega)^3},$$

and by the same arguments as above $E_2 \in H(\text{curl}, \Omega)$ with $\nabla \times E_2 = -\mu H_3 \in L^2(\Omega)^3$ and

$$\int_\Omega (\nabla \times H_1) \cdot E_2 = \int_\Omega H_1 \cdot (\nabla \times E_2) \quad \forall H_1 \in H(\text{curl}, \Omega).$$

Using Theorem 1.6 we conclude that this equation also holds for all functions $H_1 \in C^\infty(\overline{\Omega})^3$,

$$\int_\Omega (\nabla \times H_1) \cdot E_2 = \int_\Omega H_1 \cdot (\nabla \times E_2) \quad \forall H_1 \in C^\infty(\overline{\Omega})^3.$$

Lemma 1.10 then yields $E_2 \in H_0(\text{curl}, \Omega)$. Thus, we have shown $v_2 = [H_2, E_2]^T \in \mathcal{D}(A)$ and consequently the inclusion $\mathcal{D}(A^\star) \subset \mathcal{D}(A)$. □

We can draw an important consequence directly from this proposition.

Corollary 1.15 *For all $v \in \mathcal{D}(A)$ we have $(Av, v)_V = 0$.*

Now we can prove the well-posedness of the evolution equation (1.22).

Theorem 1.16 *(Well-posedness). The operator $-A$ generates a C_0-group of unitary operators*

$$T : \mathbb{R} \to \mathcal{L}(V, V), \quad t \mapsto e^{-tA}. \tag{1.24}$$

Consequently, for every initial value $u_0 \in \mathcal{D}(A)$ the homogeneous evolution equation (1.22) posses a unique solution $u \in C^1(\mathbb{R}_+; V) \cap C(\mathbb{R}_+; \mathcal{D}(A))$ given by

$$u(t) = T(t)u_0. \tag{1.25}$$

Furthermore, the electromagnetic energy is conserved,

$$\|u(t)\|_V = \|u_0\|_V \qquad \forall t \geq 0. \tag{1.26}$$

Proof: In [18, Theorem 2.2] it is proven that the homogeneous evolution equation (1.22) is well-posed if and only if the operator $-A$ generates a C_0-semigroup, say $T(\cdot)$. In this case the solution of (1.22) is given by $u = T(\cdot)u_0$, for every initial value $u_0 \in \mathcal{D}(A)$. Using Proposition 1.14 and Stone's theorem A.1 we conclude that $-A$ even generates a C_0-group of unitary operators. Conservation of the electromagnetic energy is an immediate consequence of the unitary property of the C_0-group. □

Incorporation of the div-equations Until now, we have proven the well-posedness of the abstract evolution equation (1.22) which corresponds to the reduced system of Maxwell's equations (1.7) without electric current and electric charge, $\mathcal{J} = \rho = 0$. Clearly, in this situation conservation of charge (1.2) holds true and we commented that the div-equations and the boundary condition for $\mathcal{H}$ are satisfied in this case given they hold true at an intial time. Now, we shortly illustrate how this can be mathematically formalized. First, we need the concept of the *variatonal divergence*. Motivated by the ideas concerning the variational curl we define

$$H(\mathrm{div}, \Omega) := \overline{C^\infty(\Omega)^3}, \quad H_0(\mathrm{div}, \Omega) := \overline{C_0^\infty(\Omega)^3}, \tag{1.27}$$

where the closure is taken w. r. t. the *div-norm*:

$$\|F\|^2_{H(\mathrm{div},\Omega)} := \|F\|^2_{L^2(\Omega)^3} + \|\nabla \cdot F\|^2_{L^2(\Omega)^3}.$$

Both spaces admit a variational divergence [14, Theorem 3.22]: For $F \in H(\mathrm{div}, \Omega)$ (or $F \in H_0(\mathrm{div}, \Omega)$) there is a unique $g \in L^2(\Omega)$ with

$$\int_\Omega F \cdot \nabla\varphi = -\int_\Omega g\varphi \quad \forall \varphi \in C_0^\infty(\Omega), \tag{1.28}$$

and we define $\nabla \cdot F := g$. Furthermore, it is proven in [14, Theorem 3.25] that the space $H_0(\mathrm{div}, \Omega)$ contains functions with vanishing *normal component* (compare with $H_0(\mathrm{curl}, \Omega)$, which contains functions with vanishing tangential component).
Now we derive the subspace $V_0 \subset V$ as

$$V_0 := \{[H, E]^T \in V \,|\, \mu H \in H_0(\mathrm{div}, \Omega), \varepsilon E \in H(\mathrm{div}, \Omega), \nabla \cdot (\mu H) = \nabla \cdot (\varepsilon E) = 0\} \tag{1.29}$$

and the operator A_0 on the domain $\mathcal{D}(A_0) := \mathcal{D}(A) \cap V_0$ as

$$A_0 := A|_{\mathcal{D}(A_0)}. \tag{1.30}$$

This allows us to state the whole Maxwell system (1.6) as the abstract evolution problem: Given the initial value $\widehat{u}_0 = [\widehat{\mathcal{H}}_0, \widehat{\mathcal{E}}_0]^T \in \mathcal{D}(A_0)$ we search for $\widehat{u} = [\widehat{\mathcal{H}}, \widehat{\mathcal{E}}]^T \in C^1(\mathbb{R}_+; V) \cap C(\mathbb{R}_+; \mathcal{D}(A_0))$ such that

$$\partial_t \widehat{u} + A_0 \widehat{u} = 0, \qquad t \geq 0, \tag{1.31a}$$
$$\widehat{u}(0) = \widehat{u}_0. \tag{1.31b}$$

Then, we have the following well-posedness result.

Theorem 1.17 *(Well-posedness, [10, Proposition 3.5]). The operator $-A_0$ generates a C_0-group of unitary operators*

$$T_0 : \mathbb{R} \to \mathcal{L}(V_0, V_0), \quad t \mapsto e^{-tA_0}. \tag{1.32}$$

Thus, for every $\widehat{u}_0 \in \mathcal{D}(A_0)$ the homogeneous evolution equation (1.31) has a unique solution $\widehat{u} \in C^1(\mathbb{R}_+; V_0) \cap C(\mathbb{R}_+; \mathcal{D}(A_0))$ given as

$$\widehat{u}(t) = T_0(t)\widehat{u}_0. \tag{1.33}$$

Furthermore, conservation of the electromagnetic energy holds true

$$\|\widehat{u}(t)\|_V = \|\widehat{u}_0\|_V \qquad \forall t \geq 0. \tag{1.34}$$

This theorem states that if the solution u of the homogeneous evolution problem (1.22) satisfies the div-equations and the boundary condition for $\mathcal{H}$ at time $t = 0$, i. e. the initial value u_0 belongs to $\mathcal{D}(A_0)$, then it accords with the solution $\widehat{u}$ of (1.31). This means u automatically satisfies both the div-equations and the boundary condition for $\mathcal{H}$ for every time $t > 0$.

Inhomogeneous case Finally, we consider the *inhomogeneous problem*. Therefore, we define the *source terms* $g = \widehat{g} = [0, -\varepsilon \mathcal{J}]^T$. Then, we have for the reduced system (1.7) the following abstract evolution problem: Given $u_0 = [\mathcal{H}_0, \mathcal{E}_0]^T \in \mathcal{D}(A)$ we search for $u = [\mathcal{H}, \mathcal{E}]^T \in C^1(\mathbb{R}_+; V) \cap C(\mathbb{R}_+; \mathcal{D}(A))$ such that

$$\partial_t u + Au = g, \qquad t \geq 0, \tag{1.35a}$$
$$u(0) = u_0. \tag{1.35b}$$

For the full system (1.6) the evolution problem reads as: For given $\widehat{u}_0 = [\widehat{\mathcal{H}}_0, \widehat{\mathcal{E}}_0]^T \in \mathcal{D}(A_0)$ we search for $\widehat{u} = [\widehat{\mathcal{H}}, \widehat{\mathcal{E}}]^T \in C^1(\mathbb{R}_+; V_0) \cap C(\mathbb{R}_+; \mathcal{D}(A_0))$ such that

$$\partial_t \widehat{u} + A_0 \widehat{u} = \widehat{g}, \qquad t \geq 0, \tag{1.36a}$$
$$\widehat{u}(0) = \widehat{u}_0. \tag{1.36b}$$

Using the variation of constant technique we can prove the well-posedness of these problems.

Theorem 1.18 *(Variation of Constant, [18, Theorem 2.9]).*

i) Assume that the initial value satisfes $u_0 \in \mathcal{D}(A)$. Furthermore, assume that the source term satisfies $g \in C^1(\mathbb{R}_+; V)$ or $g \in C(\mathbb{R}_+; \mathcal{D}(A))$. Then, the inhomogeneous evolution problem (1.35) has a unique soultion $u \in C^1(\mathbb{R}_+; V) \cap C(\mathbb{R}_+; \mathcal{D}(A))$ given by

$$u(t) = T(t)u_0 + \int_0^t T(t-s)g(s)\, ds. \tag{1.37}$$

ii) Assume that the initial value satisfes $\widehat{u}_0 \in \mathcal{D}(A_0)$. In addition, assume that the source term satisfies $\widehat{g} \in C^1(\mathbb{R}_+; V_0)$ or $\widehat{g} \in C(\mathbb{R}_+; \mathcal{D}(A_0))$. Then, the inhomogeneous evolution problem (1.36) has a unique soultion $\widehat{u} \in C^1(\mathbb{R}_+; V) \cap C(\mathbb{R}_+; \mathcal{D}(A_0))$ given by

$$\widehat{u}(t) = T_0(t)\widehat{u}_0 + \int_0^t T_0(t-s)\widehat{g}(s)\, ds. \tag{1.38}$$

We end this chapter by proving the stability of the solutions u and $\widehat{u}$ in the sense that they can be bounded in terms of the data, i. e. in terms of the initial values u_0, $\widehat{u}_0$ and the source terms g, $\widehat{g}$.

Theorem 1.19 *(Stability). Under the assumption of the previous Theorem 1.18 it holds*

$$\|u(t)\|_V \leq \|u_0\|_V + \int_0^t \|g(s)\|_V\, ds, \tag{1.39}$$

and

$$\|\widehat{u}(t)\|_V \leq \|\widehat{u}_0\|_V + \int_0^t \|\widehat{g}(s)\|_V\, ds, \tag{1.40}$$

Proof: This follows from (1.37) and (1.38) by using the triangle inequality and the unitary property of T and T_0. □

Chapter 2

Spatial Discretization I: The Discrete Setting

In this section we present the ingredients needed to constrcut a spatial discretization of Maxwell's equations by dG methods. Like FE methods the idea of dG methods is to construct finite dimensional function spaces in which we search for an approximate solution. The construction of this function spaces consists of two steps: First we discretize the domain Ω using a mesh and then we construct the approximation space as the space of all functions which are polynomials on each mesh element. This leads to the concept of broken polynomial spaces. Furthermore, we introduce the broken version of $H(\text{curl}, \Omega)$ which is of particular interest for the analysis of Maxwell's equations. This chapter mainly relies on the textbook [17].

2.1 Meshes

2.1.1 Basic Concepts

We make following assumption on the domain Ω.

Assumption 2.1 *(Domain Ω). The domain Ω is a polyhedron in $\mathbb{R}^d$.*

The advantage of this assumption is that polyhedra can be exactly covered by meshes built of polyhedral elements. Furthermore, it allows us to define the unit outward normal a. e.

Definition 2.2 (Boundary and outward unit normal). We denote the *boundary* of Ω by $\partial\Omega$ and the *unit outward normal* by n.

We start introducing meshes with the simple case of simplicial meshes.

Definition 2.3 (Simplex). Let $\{x_0, \dots, x_d\}$ be a set of $d+1$ points in $\mathbb{R}^d$ such that the vectors $\{x_1 - x_0, \dots, x_d - x_0\}$ are linearly independent. Then, the interior of the convex hull of $\{x_0, \dots, x_d\}$ is called a non-degenerate *simplex* of $\mathbb{R}^d$, and the points $\{x_0, \dots, x_d\}$ its *vertices*.

In dimension 1, a non-degenerate simplex is an open interval, in dimension 2 a triangle and in dimension 3 a tetrahedron.

Definition 2.4 (Simplicial mesh). A finite set $\mathcal{T} = \{K\}$ is called a *simplicial mesh* of the domain Ω if:

i) Every $K \in \mathcal{T}$ is a non-degenerate simplex,

ii) the set $\mathcal{T}$ forms a partition of Ω, i. e.

$$\overline{\Omega} = \bigcup_{K \in \mathcal{T}} \overline{K},$$

and for every $K_1, K_2 \in \mathcal{T}$, $K_1 \neq K_2$, it holds

$$K_1 \cap K_2 = \emptyset.$$

Each $K \in \mathcal{T}$ is called a *mesh element*.

When working with finite elements simplicial meshes are a quite convenient choice. An advantage of dG methods is that they allow working with general meshes more easily.

Definition 2.5 (General mesh). A *general mesh* $\mathcal{T}$ of the domain Ω is a finite collection of polyhedra $\mathcal{T} = \{K\}$ satisfying condition ii) of the previous Definition 2.4. Each element $K \in \mathcal{T}$ is called a *mesh element*.

Obviously, a simplicial mesh is a particular case of a general mesh.

Definition 2.6 (Element diameter, meshsize). Let $\mathcal{T}$ be a general mesh of the domain Ω. We denote with h_K the *diameter* of a mesh element $K \in \mathcal{T}$. Furthermore, we define the *meshsize* h as the largerst diameter in the mesh

$$h := \max_{K \in \mathcal{T}} h_K.$$

In what follows we will refer to a mesh $\mathcal{T}$ with meshsize h with $\mathcal{T}_h$.

Definition 2.7 (Element outward normal). Let $\mathcal{T}_h$ be a mesh of the domain Ω and $K \in \mathcal{T}_h$. We define n_K a. e. on ∂K as the unit outward normal to K.

2.1.2 Mesh Faces, Averages and Jumps

Now we introduce the concept of mesh faces, averages and jumps, which will be frequently used in the design and analysis of dG methods.

Definition 2.8 (Mesh faces). Let $\mathcal{T}_h$ be a mesh of the domain Ω. We say that a closed subset F of $\overline{\Omega}$ is a *mesh face* if F has positive $(d-1)$-measure and either one of the two following conditions is satisfied:

i) There are distinct mesh elements $K_1, K_2 \in \mathcal{T}_h$ such that $F = \partial K_1 \cap \partial K_2$; in this case, we call F an *interface*.

ii) There is a mesh element $K \in \mathcal{T}_h$ such that $F = \partial K \cap \partial\Omega$; in this case, we call F a *boundary face*.

We collect interfaces in the set $\mathcal{F}_h^i$ and boundary faces in the set $\mathcal{F}_h^b$. Henceforth, we set

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

Furthermore, for any mesh element $K \in \mathcal{T}_h$ we collect the mesh faces composing the boundary of K in the set

$$\mathcal{F}_K := \{F \in \mathcal{F}_h \,|\, F \subset \partial K\}.$$

Finally, we denote the maximum number of mesh faces composing the boundary of mesh elements by

$$N_\partial := \max_{K \in \mathcal{T}_h} \mathrm{card}(\mathcal{F}_K).$$

Figure 2.1 illustrates four interfaces and one boundary face of a mesh element belonging to a general mesh.

We continue with the definition of face normals. Therefore, we introduce the following notation which we will keep from now on: For every mesh element $K \in \mathcal{T}_h$ and every corresponding interface $F \in \mathcal{F}_K \cap \mathcal{F}_h^i$ we denote the neighboring mesh element w. r. t. F with K_F, see Figure 2.2.

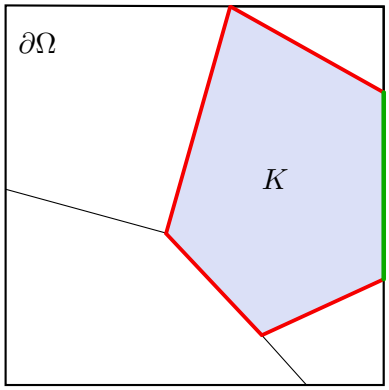


Figure 2.1: Example of interfaces (*red*) and boundary face (*green*)

Definition 2.9 (Face normals). For all $F \in \mathcal{F}_h$ we define the *unit normal* n_F to F as

i) the unit normal n_K to F pointing from K to K_F if $F \in \mathcal{F}_h^i$, see Figure 2.2; the orientation of n_F is arbitrary depending on the choice of K, but kept fixed in what follows.

ii) The outward unit normal n to Ω if $F \in \mathcal{F}_h^b$.

Next, we turn to averages and jumps across interfaces of piecewise smooth functions. Let us therefore introduce the following notation

$$v_K := v|_K, \quad v_{K_F} := v|_{K_F}.$$

Definition 2.10 (Interface averages and jumps). Let v be a scalar-valued function and assume that for every mesh element $K \in \mathcal{T}_h$ its restricition $v|_K$ is smooth enough to admit a trace a. e. on the boundary ∂K. Then, for all $F \in \mathcal{F}_h^i$ the function v admits a possible two-valued trace and we define

i) the *average* of v on F as

$$\{\{v\}\}_F := \frac{1}{2}\big((v_K)|_F + (v_{K_F})|_F\big),$$

ii) the *jump* of v on F as

$$[\![v]\!]_F := (v_{K_F})|_F - (v_K)|_F.$$

When v is vector-valued, the above average and jump operators act componentwise on v.

2.1.3 Broken Polynomial Spaces

By now we have constructed a mesh of the domain Ω and so we can turn to the second step, namely to the construction of finite function spaces. In our case this spaces consist of piecewise polynomials.

The polynomial space $\mathbb{P}_d^k$

Let $k \geq 0$ be an integer and $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$ be a multi-index with

$$|\alpha|_{l^1} = \sum_{i=1}^{d} \alpha_i \leq k.$$

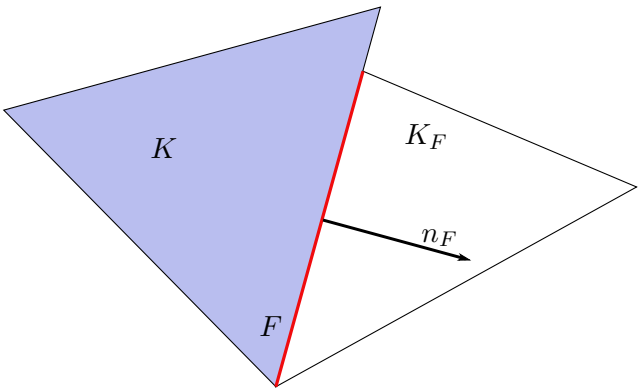


Figure 2.2: Basic notation

Further let $x = (x_1, \dots, x_d)$ be a vector in $\mathbb{R}^d$ and let us use the convention

$$x^\alpha := \prod_{i=1}^{d} x_i^{\alpha_i}.$$

Then, the function p_α defined as

$$p_\alpha : \mathbb{R}^d \to \mathbb{R}, \quad x \mapsto \gamma_\alpha x^\alpha,$$

where $\gamma_\alpha \in \mathbb{R}$ is a coefficient, is a polynomial of d variables of total degree at most k. Consequently, the set

$$\mathbb{P}_d^k := \left\{ p : \mathbb{R}^d \to \mathbb{R} \,\middle|\, \exists(\gamma_\alpha) \subset \mathbb{R} \text{ s.t. } p(x) = \sum_{|\alpha|_l{}^1 \le k} \gamma_\alpha x^\alpha \right\}$$

is the space of all polynomials of d variables with degree at most k. Its dimension is

$$\dim(\mathbb{P}_d^k) = \binom{k+d}{k} = \frac{(k+d)!}{k!d!}.$$

The broken polynomial space $\mathbb{P}_d^k(\mathcal{T}_h)$

Let $K \in \mathcal{T}_h$ be a mesh element. Then, we define $\mathbb{P}_d^k(K)$ as

$$\mathbb{P}_d^k(K) := \left\{ p|_K : K \to \mathbb{R} \,\middle|\, p \in \mathbb{P}_d^k \right\}.$$

The *broken polynomial space* $\mathbb{P}_d^k(\mathcal{T}_h)$ on the mesh $\mathcal{T}_h$ now consists of functions which are polynomials on each mesh element, i. e.

$$\mathbb{P}_d^k(\mathcal{T}_h) := \left\{ v \in L^2(\Omega) \,|\, \forall K \in \mathcal{T}_h : v|_K \in \mathbb{P}_d^k(K) \right\}. \tag{2.1}$$

It follows that

$$\dim(\mathbb{P}_d^k(\mathcal{T}_h)) = \text{card}(\mathcal{T}_h) \times \dim(\mathbb{P}_d^k).$$

2.1.4 Broken Sobolev Spaces

After having introduced polynomial spaces and their broken versions, we now consider broken versions of Sobolev spaces $H^m(\Omega)$ and of the graph space $H(\text{curl}, \Omega)$ as well as broken versions of the gradient and the curl operator.

The broken Sobolev space $H^m(\mathcal{T}_h)$

Let $m \geq 0$ be an integer. We define the *broken Sobolev space* as

$$H^m(\mathcal{T}_h) := \left\{ v \in L^2(\Omega) \,|\, \forall K \in \mathcal{T}_h : v|_K \in H^m(K) \right\}, \tag{2.2}$$

and endow it with the norm: For $v \in H^m(\mathcal{T}_h)$,

$$\|v\|^2_{H^m(\mathcal{T}_h)} := \sum_{n=0}^{m} |v|^2_{H^n(\mathcal{T}_h)}, \qquad |v|_{H^n(\mathcal{T}_h)} := \sum_{K \in \mathcal{T}_h} |v|^2_{H^n(K)}.$$

Using the continuous trace inequality [17, Chapter 1] we see that for all functions $v \in H^1(\mathcal{T}_h)$ and for all mesh elements $K \in \mathcal{T}_h$ the trace $v|_{\partial K}$ on the boundary of the element is well-defined and it holds

$$\|v\|_{L^2(\partial K)} \leq C \|v\|^{1/2}_{L^2(K)} \|v\|^{1/2}_{H^1(K)}.$$

It is natural to define a broken gradient operator acting on the broken Sobolev space $H^1(\mathcal{T}_h)$. Clearly, this operator then also acts on the broken polynomial spaces.

Definition 2.11 (Broken gradient). The *broken gradient* $\nabla_h : H^1(\mathcal{T}_h) \to L^2(\Omega)^d$ is defined such that, for all $v \in H^1(\mathcal{T}_h)$,

$$(\nabla_h v)|_K := \nabla(v|_K), \quad \forall K \in \mathcal{T}_h. \tag{2.3}$$

It is important to distinguish the usual Sobolev spaces from their broken versions and we now characterize this in more detail. Clearly, the usual Sobolev spaces are subspaces of their broken versions, i. e. for every integer $m \geq 0$, we have

$$H^m(\Omega) \subset H^m(\mathcal{T}_h).$$

Furthermore, it is proven in [17, Lemma 1.22] that for functions in $H^1(\Omega)$ the (variational) gradient coincides with the broken gradient: For all $v \in H^1(\Omega)$,

$$\nabla v = \nabla_h v.$$

But, in general the reverse does not hold true. The crucial difference is that the broken Sobolev spaces contain functions having nonzero jumps at interfaces whereas functions in the usual Sobolev spaces must have zero jumps across any interface. The exact statement reads as follows.

Lemma 2.12 *(Charaterization of $H^1(\Omega)$, [17, Lemma 1.23]). A function $v \in H^1(\mathcal{T}_h)$ belongs to $H^1(\Omega)$ if and only if*

$$[\![v]\!]_F = 0 \qquad \forall F \in \mathcal{F}_h^i. \tag{2.4}$$

The broken graph space $H(\mathrm{curl}, \mathcal{T}_h)$

Analogously to the definition of the broken Sobolev spaces $H^m(\mathcal{T}_h)$ we introduce the broken version of the graph space $H(\mathrm{curl}, \Omega)$ as

$$H(\mathrm{curl}, \mathcal{T}_h) := \left\{ v \in L^2(\Omega)^3 \,|\, \forall K \in \mathcal{T}_h : v \in H(\mathrm{curl}, K) \right\}. \tag{2.5}$$

Naturally, we also introduce a broken version of the curl operator.

Definition 2.13 (Broken curl). The *broken curl* $\nabla_h \times : H(\mathrm{curl}, \mathcal{T}_h) \to L^2(\Omega)^3$ is defined such that, for all $v \in H(\mathrm{curl}, \mathcal{T}_h)$,

$$(\nabla_h \times v)|_K := \nabla \times (v|_K) \qquad \forall K \in \mathcal{T}_h. \tag{2.6}$$

The relation between $H(\mathrm{curl}, \Omega)$ and its broken version $H(\mathrm{curl}, \mathcal{T}_h)$ is featured by a similiar result as for the previously considered Sobolev spaces and their broken counterparts.

Lemma 2.14 *(Broken curl on $H(\mathrm{curl}, \Omega)$). There holds $H(\mathrm{curl}, \Omega) \subset H(\mathrm{curl}, \mathcal{T}_h)$. Moreover, for all $v \in H(\mathrm{curl}, \Omega)$,*

$$\nabla_h \times v = \nabla \times v. \tag{2.7}$$

Proof: We adapt the proof of the inclusion $H^1(\Omega) \subset H^1(\mathcal{T}_h)$ given in [17, Lemma 1.22]. Let $v \in H(\mathrm{curl}, \Omega)$ and $K \in \mathcal{T}_h$. Further let $\varphi \in C_0^\infty(K)^3$ and let $E\varphi$ denote its extension by zero to Ω. Clearly, we have $E\varphi \in C_0^\infty(\Omega)^3$. Then, it holds

$$\int_K v|_K \cdot (\nabla \times \varphi) = \int_\Omega v \cdot (\nabla \times E\varphi) = \int_\Omega (\nabla \times v) \cdot E\varphi = \int_K (\nabla \times v)|_K \cdot \varphi,$$

where the second equality holds true by the definition of $H(\mathrm{curl}, \Omega)$ (1.14) and the fact that $E\varphi$ is in $C_0^\infty(\Omega)^3$. The same definition, now applied on K, yields that $v|_K$ is an element of $H(\mathrm{curl}, K)$ and that it holds

$$\nabla \times (v|_K) = (\nabla \times v)|_K.$$

Since the element K was arbitrary this holds true for all $K \in \mathcal{T}_h$. Recalling Definition 2.13 of the broken curl it follows

$$(\nabla_h \times v)|_K = (\nabla \times v)|_K \qquad \forall K \in \mathcal{T}_h,$$

which is the stated claim. □

In the later work we will often consider the space $H(\mathrm{curl}, \Omega) \cap H^1(\mathcal{T}_h)^3$ and it turns out to be crucial that its functions only admit *zero tangential jumps* across interfaces.

Lemma 2.15 *(Characterization of $H(\mathrm{curl}, \Omega)$). A function $v \in H^1(\mathcal{T}_h)^3$ belongs to $H(\mathrm{curl}, \Omega)$ if and only if*

$$n_F \times [\![v]\!]_F = 0 \qquad \forall F \in \mathcal{F}_h^i. \tag{2.8}$$

Proof: We follow the proof given in [16, Lemma 3.4]. Let $v \in H^1(\mathcal{T}_h)^3$. We start by proving that condition (2.8) is sufficient. Thus, assume that it holds $n_F \times [\![v]\!]_F = 0$ for all $F \in \mathcal{F}_h^i$. Then, for any $\varphi \in C_0^\infty(\Omega)^3$, it holds

$$\int_\Omega (\nabla_h \times v) \cdot \varphi = \sum_{K \in \mathcal{T}_h} \int_K (\nabla \times v|_K) \cdot \varphi = \sum_{K \in \mathcal{T}_h} \left(\int_K v \cdot (\nabla \times \varphi) + \int_{\partial K} (n_K \times v) \cdot \varphi \right). \tag{2.9}$$

Here we used that $H^1(\mathcal{T}_h)^3$ is a subspace of $H(\mathrm{curl}, \mathcal{T}_h)$ and thus the broken curl is well-defined. Furthermore, we used that for every mesh element K the function v is in $H^1(K)^3$ and thus partial integration is applicable. Now let us consider the sum over the boundaries ∂K in the upper equation. Owing to the convention about face normals (see Definition 2.9) and since φ vanishes on the boundary faces, it holds

$$\begin{aligned} \sum_{K \in \mathcal{T}_h} \int_{\partial K} (n_K \times v) \cdot \varphi &= \sum_{F \in \mathcal{F}_h^i} \left(\int_F (n_K \times v_K) \cdot \varphi + \int_F (n_{K_F} \times v_{K_F}) \cdot \varphi \right) \\ &\quad + \sum_{F \in \mathcal{F}_h^b} \int_F (n \times v) \cdot \varphi \\ &= - \sum_{F \in \mathcal{F}_h^i} \int_F (n_F \times [\![v]\!]_F) \cdot \varphi. \end{aligned} \tag{2.10}$$

Since we assumed that the tangential jumps of v vanish we have

$$\int_\Omega (\nabla_h \times v) \cdot \varphi = \sum_{K \in \mathcal{T}_h} \int_K v \cdot (\nabla \times \varphi) = \int_\Omega v \cdot (\nabla \times \varphi).$$

As this holds true for all $\varphi \in C_0^\infty(\Omega)^3$ we can conlcude with (1.14) that $v \in H(\mathrm{curl}, \Omega)$ and that $\nabla \times v = \nabla_h \times v$. Thus, we have shown sufficiency.
Now let us prove that $v \in H(\mathrm{curl}, \Omega)$ is a necessary condition for (2.8) to hold. So let $v \in H^1(\mathcal{T}_h)^3 \cap H(\mathrm{curl}, \Omega)$. Then, using equations (2.9) and (2.10), we infer for all $\varphi \in C_0^\infty(\Omega)^3$

$$\int_\Omega (\nabla_h \times v) \cdot \varphi = \int_\Omega v \cdot (\nabla \times \varphi) - \sum_{F \in \mathcal{F}_h^i} \int_F (n_F \times [\![v]\!]_F) \cdot \varphi. \tag{2.11}$$

On the other hand, we see by Lemma 2.14 and Definition 1.2, that for all $\varphi \in C_0^\infty(\Omega)^3$ there holds

$$\int_\Omega (\nabla_h \times v) \cdot \varphi = \int_\Omega (\nabla \times v) \cdot \varphi = \int_\Omega v \cdot (\nabla \times \varphi). \tag{2.12}$$

Combining (2.11) and (2.12) yields for all $\varphi \in C_0^\infty(\Omega)^3$

$$\sum_{F \in \mathcal{F}_h^i} \int_F (n_F \times [\![v]\!]_F) \cdot \varphi = 0.$$

Hence, (2.8) follows by choosing the support of φ intersecting a single interface and since φ is arbitrary. □

2.2 Admissible Mesh Sequences

The last section in this chapter is dedicated to the concept of admissible mesh sequences. This is of special interest since we would like to prove convergence of dG methods, which means that the error between the approximate solution and the exact solution tends to zero as the meshsize goes to zero. So let us consider the mesh sequence

$$\mathcal{T}_\mathcal{H} := (\mathcal{T}_h)_{h \in \mathcal{H}},$$

where $\mathcal{H}$ denotes a countable subset $\mathbb{R}_+$ having 0 as only accumulation point. In the following we consider so called shape- and contact-regular mesh sequences, see [17, Definition 1.38], which posses the properties we need for the convergence analysis. We only state these properties and refer to [17, Section 1.4.1] for a detailed insight into this topic.

2.2.1 Geometric Properties

We just need one geometric property. Recall that we have defined $\mathcal{F}_K$ as the set of faces composing the boundary of an element K and N_∂ as the maximum number of mesh faces composing the boundary of elements in $\mathcal{T}_h$, see Section 2.1.2. Then, for shape- and contact-regular meshes, we can characterize the relation between these quantities and the meshsize h by following lemma.

Lemma 2.16 *(Bound on card$(\mathcal{F}_K)$ and N_∂, [17, Lemma 1.41]). Let $\mathcal{T}_\mathcal{H}$ be a shape- and contact-regular mesh sequence. Then, for all $h \in \mathcal{H}$ and all $K \in \mathcal{T}_h$, card$(\mathcal{F}_K)$ and N_∂ are uniformly bounded in h.*

2.2.2 Inverse and Trace Inequality

We proceed by stating two inequalities for the broken polynomial spaces $\mathbb{P}_d^k(\mathcal{T}_h)$ on a shape- and contact-regular mesh. This inequalities turn out to be very useful for analyzing dG methods. The inverse inequality provides an upper bound on the gradient of discrete functions.

Lemma 2.17 *(Inverse inequality, [17, Lemma 1.44]). Let $\mathcal{T}_\mathcal{H}$ be a shape- and contact-regular mesh sequence. Then, for all $h \in \mathcal{H}$, all $K \in \mathcal{T}_h$ and all $v_h \in \mathbb{P}_d^k(\mathcal{T}_h)$,*

$$\|\nabla v_h\|_{L^2(K)^d} \leq C_{\mathrm{inv}} h_K^{-1} \|v_h\|_{L^2(K)}. \tag{2.13}$$

The constant C_{inv} only depends on d, k and the shape- and contact-regularity parameters.

The second inequality is a discrete trace inequality that delivers an upper bound on the face values of discrete functions.

Lemma 2.18 *(Discrete trace inequality, [17, Lemma 1.46]). Let $\mathcal{T}_\mathcal{H}$ be a shape- and contact-regular mesh sequence. Then, for all $h \in \mathcal{H}$, all $K \in \mathcal{T}_h$, all $F \in \mathcal{F}_K$ and all $v_h \in \mathbb{P}_d^k(\mathcal{T}_h)$,*

$$\|v_h\|_{L^2(F)} \le C_{\mathrm{tr}} h_K^{-1/2} \|v_h\|_{L^2(K)}. \tag{2.14}$$

The constant C_{tr} only depends on d, k and the shape- and contact-regularity parameters.

Remark 2.19 (k-dependency). The constant C_{inv} scales as k^2 (on triangles), whereas the constant C_{tr} scales as $\sqrt{k(k+d)}$, see [17, Remark 1.48]. ◇

Lemma 2.20 *(Continuous trace inequality, [17, Lemma 1.49]) Let $\mathcal{T}_\mathcal{H}$ be a shape- and contact-regular mesh sequence. Then, for all $h \in \mathcal{H}$, all $K \in \mathcal{T}_h$, all $F \in \mathcal{F}_K$ and all $v \in H^1(\mathcal{T}_h)$,*

$$\|v\|_{L^2(F)}^2 \le C_{\mathrm{cti}}(2\|\nabla v\|_{L^2(K)^d} + d h_K^{-1} \|v\|_{L^2(K)}) \|v\|_{L^2(K)}, \tag{2.15}$$

where the constant C_{cti} depends on d and the shape- and contact-regularity parameters.

2.2.3 Polynomial Approximation

In dG methods we search the approximate solution in the piecewise polynomial space $\mathbb{P}_d^k(\mathcal{T}_h)$. Thus, it is important to ensure that the mesh sequence is constructed such that optimal polynomial approximation hold true. In order to include this consideration we require that the mesh sequence admits optimal polynomial approximation in the sense of the next definition. Again we refer to [17, Section 1.4.4] for details.

Definition 2.21 (Optimal polynomial approximation). The mesh sequence $\mathcal{T}_\mathcal{H}$ has *optimal polynomial approximation* properties if, for all $h \in \mathcal{H}$, all $K \in \mathcal{T}_h$, and all polynomial degree k, there is a linear interpolation operator $\mathcal{I}_K^k : L^2(K) \to \mathbb{P}_d^k(K)$ such that, for all $s \in \{0, \dots, k+1\}$ and all $v \in H^s(K)$, there holds

$$|v - \mathcal{I}_K^k v|_{H^m(K)} \le C_{\mathrm{app}} h_K^{s-m} |v|_{H^s(K)} \qquad \forall m \in \{0, \dots, s\}, \tag{2.16}$$

where the constant C_{app} is independent of both K and h.

Now we have introduced all ingredients and it remains to built an admissible mesh sequence.

Assumption 2.22 *(Admissible mesh sequence). We assume that the mesh-sequence $\mathcal{T}_\mathcal{H}$ is* admissible*, i. e. that it is shape- and contact-regular and has optimal polynomial approximation properties in the sense of Definition 2.21.*

In the later error analysis we will often use the L^2*-orthogonal projection* onto the broken polynomial space $\mathbb{P}_d^k(\mathcal{T}_h)$ defined as, $\pi_h^{L^2} : L^2(\Omega) \to \mathbb{P}_d^k(\mathcal{T}_h)$ such that for all $v \in L^2(\Omega)$,

$$(\pi_h^{L^2} v, \varphi_h)_{L^2(\Omega)} = (v, \varphi_h)_{L^2(\Omega)} \qquad \forall \varphi_h \in \mathbb{P}_d^k(\mathcal{T}_h). \tag{2.17}$$

Admissible mesh sequences provide optimality of the L^2-projetion in the following sense.

Lemma 2.23 *(Optimality of L^2-orthogonal projection). Let $\mathcal{T}_\mathcal{H}$ be an admissible mesh sequence. Then, for all $s \in \{0, \dots, k+1\}$ and all $v \in H^s(K)$, there holds*

$$|v - \pi_h^{L^2} v|_{H^m(K)} \le C'_{\mathrm{app}} h_K^{s-m} |v|_{H^s(K)} \qquad \forall m \in \{0, \dots, s\}. \tag{2.18}$$

The constant C'_{app} is independent of both K and h.

Proof: We follow the proof in [17, Lemma 1.58]. Let $v \in H^s(K)$. We first consider the case $m = 0$. Clearly, the projection and the interpolation of v are functions in $\mathbb{P}_d^k(\Omega)$, i. e. $\pi_h^{L^2} v$, $\mathcal{I}_K^k v \in \mathbb{P}_d^k(\Omega)$. Owing to the definition of the projection (2.17) this yields

$$\begin{aligned} 0 =& (v - \pi_h^{L^2} v, \mathcal{I}_K^k v - \pi_h^{L^2} v)_{L^2(K)} = (v - \pi_h^{L^2} v, \mathcal{I}_K^k v - v + v - \pi_h^{L^2} v)_{L^2(K)} \\ =& (v - \pi_h^{L^2} v, \mathcal{I}_K^k v - v)_{L^2(K)} + \|v - \pi_h^{L^2} v\|_{L^2(K)}. \end{aligned}$$

Applying the Cauchy-Schwarz inequality we get

$$\|v - \pi_h^{L^2} v\|_{L^2(K)} \leq \|v - \mathcal{I}_K^k v\|_{L^2(K)} \leq C_{\mathrm{app}} h_K^s |v|_{H^s(K)}, \tag{2.19}$$

where the second inequality is due to (2.16). This concludes the case $m = 0$ and we proceed with $m \geq 1$. We use m times the inverse inequality (2.13), the triangle inequality and (2.19) to infer

$$\begin{aligned} |v - \pi_h^{L^2} v|_{H^m(K)} \leq& |v - \mathcal{I}_K^k v|_{H^m(K)} + |\mathcal{I}_K^k v - \pi_h^{L^2} v|_{H^m(K)} \\ \leq& |v - \mathcal{I}_K^k v|_{H^m(K)} + C h_K^{-m} \|\mathcal{I}_K^k v - \pi_h^{L^2} v\|_{L^2(K)} \\ =& |v - \mathcal{I}_K^k v|_{H^m(K)} + C h_K^{-m} \|\mathcal{I}_K^k v - v + v - \pi_h^{L^2} v\|_{L^2(K)} \\ \leq& |v - \mathcal{I}_K^k v|_{H^m(K)} + 2C h_K^{-m} \|v - \mathcal{I}_K^k v\|_{<L^2(K)}. \end{aligned}$$

Now the result follows with (2.16). □

A direct consequence of (2.18) and the continuous trace inequality from Lemma 2.20 is the following bound for polynomial approximations on mesh faces.

Lemma 2.24 *(Polynomial approximation on mesh faces). Under the assumption of Lemma 2.23 with $s \geq 1$ it holds for all $h \in \mathcal{H}$, all $K \in \mathcal{T}_h$, and all $F \in \mathcal{F}_K$,*

$$\|v - \pi_h^{L^2} v\|_{L^2(F)} \leq C''_{\mathrm{app}} h_K^{s-1/2} |v|_{H^s(K)}, \tag{2.20}$$

with constant C''_{app} independent of both K and h.

Chapter 3

Spatial Discretization II: Discretization of Maxwell's Equations

In Chapter 2 we have introduced the main ingredients of a dG discretization. In this chapter we discretize Maxwell's equations in space by proceeding in following steps:

First we consider the case of homogeneous media where the coefficients ε and μ are constant. This enables us to rewrite Maxwell's equations in the normalized form as used for example in [2]. We first consider homogeneous media since it allows us to nicely illustrate the construction of the dG discretization. We begin by deriving a basic, so called centered fluxes scheme. The concept of local residuals, see e.g. [3], then motivates the idea of adding a stabilization to the centered fluxes scheme yielding an improved method called an upwind fluxes scheme.

Then, we carry over the developped ideas to the more general case of composite media consisting of different materials. This case allows piecewise constant coefficients ε and μ and therefore we need to adapt the derived schemes. We do this with the concept of local impedance and conductivity reflecting the physics of such media. This provides us with the centered fluxes scheme. The stabilized upwind fluxes scheme is again obtained via the analysis of local residuals. Another way of deriving this dG discretizations can be found in the textbook [8], where the upwind fluxes scheme is constructed via Riemann solvers.

Finally, we prove stability of both dG discretizations and show that the discretization error is featured by the same stability result. This allows us to prove the convergence of order h^k. We end the chapter by improving the convergence result to $h^{k+1/2}$ for upwind fluxes relying on different arguments than the stability result.

3.1 Homogeneous Medium

We assume for this section that the coefficients ε and μ are positive constants. Let us begin by shortly revisiting the considerations from Chapter 1. There we have introduced the state space V and the graph space $\mathcal{D}(A)$ as

$$V = L^2(\Omega)^3 \times L^2(\Omega)^3, \qquad \mathcal{D}(A) = H(\mathrm{curl}, \Omega) \times H_0(\mathrm{curl}, \Omega),$$

and stated Maxwell's equations as the following evolution problem (see (1.35)): Given an initial value $u_0 = [\mathcal{H}_0, \mathcal{E}_0]^T \in \mathcal{D}(A)$ we search for $u = [\mathcal{H}, \mathcal{E}]^T \in C^1(0,T;V) \cap C(0,T;\mathcal{D}(A))$ with $u(0) = u_0$ and such that

$$\partial_t \mathcal{H} + \mu^{-1} \nabla \times \mathcal{E} = 0 \qquad \text{in } (0,T) \times \Omega, \tag{3.1a}$$

$$\partial_t \mathcal{E} - \varepsilon^{-1} \nabla \times \mathcal{H} = -\varepsilon^{-1} \mathcal{J} \qquad \text{in } (0,T) \times \Omega. \tag{3.1b}$$

Note that from now on we restrict our considerations to bounded time intervals.

3.1.1 Normalized Form

Since the coefficients ε and μ are constant we can rewrite (3.1) as

$$\partial_t \widetilde{\mathcal{H}} + c_0 \nabla \times \widetilde{\mathcal{E}} = 0 \qquad \text{in } (0,T)\times\Omega, \tag{3.2a}$$
$$\partial_t \widetilde{\mathcal{E}} - c_0 \nabla \times \widetilde{\mathcal{H}} = -\widetilde{\mathcal{J}} \qquad \text{in } (0,T)\times\Omega, \tag{3.2b}$$

where $c_0 := (\varepsilon\mu)^{-1/2}$ is the speed of light in the medium and we set

$$\widetilde{\mathcal{H}} := \mu^{1/2}\mathcal{H}, \quad \widetilde{\mathcal{E}} := \varepsilon^{1/2}\mathcal{E}, \quad \widetilde{\mathcal{J}} := \varepsilon^{-1/2}\mathcal{J}.$$

In this formulation all appearing quantities are normalized to the same physical unit. The space discretization is based on the following equivalent formulation of (3.2): Given $\widetilde{u}_0 = [\widetilde{\mathcal{H}}_0, \widetilde{\mathcal{E}}_0]^T \in \mathcal{D}(A)$ we seek for $\widetilde{u} = [\widetilde{\mathcal{H}}, \widetilde{\mathcal{E}}]^T \in C^1(0,T;V) \cap C(0,T;\mathcal{D}(A))$ such that for all *test functions* in the state space $\varphi = [\phi, \psi]^T \in V$ it holds

$$\begin{aligned}(\partial_t \widetilde{\mathcal{H}}, \phi)_{L^2(\Omega)^3} + c_0(\nabla \times \widetilde{\mathcal{E}}, \phi)_{L^2(\Omega)^3} \\ + (\partial_t \widetilde{\mathcal{E}}, \psi)_{L^2(\Omega)^3} - c_0(\nabla \times \widetilde{\mathcal{H}}, \psi)_{L^2(\Omega)^3} = (-\widetilde{\mathcal{J}}, \psi)_{L^2(\Omega)^3}.\end{aligned} \tag{3.3}$$

We collect the appearing inner products in two bilinear forms depending on the kind of derivative they involve.

Definition 3.1 (Continuous bilinear forms). We define the bilinear forms $\widetilde{m}$, $\widetilde{a} : \mathcal{D}(A) \times V \to \mathbb{R}$ as follows: For $v = [H, E]^T$ and $\varphi = [\phi, \psi]^T$,

$$\begin{aligned}\widetilde{m}(v,\varphi) &:= (H,\phi)_{L^2(\Omega)^3} + (E,\psi)_{L^2(\Omega)^3}, \\ \widetilde{a}(v,\varphi) &:= c_0(\nabla \times E, \phi)_{L^2(\Omega)^3} - c_0(\nabla \times H, \psi)_{L^2(\Omega)^3}.\end{aligned}$$

With this notation we can write (3.3) shortly as: We search for $\widetilde{u} \in C^1(0,T;V) \cap C(0,T;\mathcal{D}(A))$ such that

$$\widetilde{m}(\partial_t \widetilde{u}, \varphi) + \widetilde{a}(\widetilde{u}, \varphi) = (\widetilde{g}, \varphi)_{L^2(\Omega)^6} \qquad \forall \varphi \in V, \tag{3.4}$$

where we set $\widetilde{g} := [0, -\widetilde{\mathcal{J}}]^T$. The dG discretization consists in replacing the continuous bilinear forms by discretized ones. This allows us to approximate the continuous problem (3.4) in a finite dimensional space making it accessible for solving on computers. The first bilinear form, $\widetilde{m}$, can be easily handled since it involves only derivatives w. r. t. to the time variable t. So our focus lies nearly solely on the discretization of the bilinear form $\widetilde{a}$.

3.1.2 Discrete Bilinear Forms

The dG discretization works with discontinuous, elementwise smooth functions and we will frequently have to consider averages and jumps over interfaces of such functions. Since functions in the graph space $\mathcal{D}(A)$ do not necessarily admit an L^2-trace we shall require slightly more regularity from the exact solution.

Assumption 3.2 *(Regularity of exact solution and space $V_\star$). We assume that the exact solution $\widetilde{u} = [\mathcal{H}, \mathcal{E}]^T$ of (3.4) satisfies*

$$\widetilde{u} \in V_\star := \mathcal{D}(A) \cap (H^1(\mathcal{T}_h)^3 \times H^1(\mathcal{T}_h)^3). \tag{3.5}$$

Let us shortly explain the consequences of this assumption: Recalling Remark 1.11 we see that the $\widetilde{\mathcal{E}}$-field vanishes on boundary faces, i. e.

$$n \times \widetilde{\mathcal{E}} = 0 \qquad \forall F \in \mathcal{F}_h^b. \tag{3.6}$$

Furthermore, we see with Lemma 2.15 that the exact solution does only admit zero tangential jumps on interfaces,

$$n_F \times [\![\widetilde{\mathcal{H}}]\!]_F = n_F \times [\![\widetilde{\mathcal{E}}]\!]_F = 0 \qquad \forall F \in \mathcal{F}_h^i. \tag{3.7}$$

We continue by introducing two more spaces needed to construct the discrete bilinear forms. We want to construct the discrete solution in the broken polynomial space $\mathbb{P}_3^k(\mathcal{T}_h)^3 \times \mathbb{P}_3^k(\mathcal{T}_h)^3$ defined in (2.1), assuming that $\mathcal{T}_h$ belongs to an admissible mesh sequence. Consequently, we define the *discrete solution space* as

$$V_h := \mathbb{P}_3^k(\mathcal{T}_h)^3 \times \mathbb{P}_3^k(\mathcal{T}_h)^3.$$

Note that the discrete solution space is not contained in the continuous solution space, $V_h \nsubseteq V_\star$ (see e.g. Lemma 2.15), which characterizes dG methods as *non-conforming* methods. Therefore, we additionally consider the space

$$V_{\star h} := V_\star + V_h,$$

which contains both the exact and the discrete solutions. In particular, $V_{\star h}$ also contains the error function of the discretization, which is just the difference of the exact and the discrete solution. Ensuring that the error function can be plugged into the first argument of the discrete bilinear forms is crucial for the later convergence analysis.

The bilinear form $\widetilde{m}_h$

The discrete version of the bilinear form $\widetilde{m}$ is obtained by defining $\widetilde{m}_h : V_{\star h} \times V_h \to \mathbb{R}$ such that for $v = [H, E]^T$ and $\varphi_h = [\phi_h, \psi_h]^T$,

$$\widetilde{m}_h(v, \varphi_h) := (H, \phi_h)_{L^2(\Omega)^3} + (E, \psi_h)_{L^2(\Omega)^3}. \tag{3.8}$$

The basic bilinear form $\widetilde{a}_h^{\mathrm{cf}}$

The ansatz to construct a discrete version of $\widetilde{a}$ consists in just replacing the curl operator by its broken version. So let $a_h^{(0)} : V_{\star h} \times V_h \to \mathbb{R}$ such that for $v = [H, E]^T$ and $\varphi_h = [\phi_h, \psi_h]^T$,

$$\widetilde{a}_h^{(0)}(v, \varphi_h) := c_0(\nabla_h \times E, \phi_h)_{L^2(\Omega)^3} - c_0(\nabla_h \times H, \psi_h)_{L^2(\Omega)^3}.$$

Obviously, the question wether we have chosen a meaningful discrete version of $\widetilde{a}$ arises. We can approach this question by checking if $\widetilde{a}_h^{(0)}$ satisfies two basic features. The first one is *consistency*, i. e. if the discrete bilinear form coincides with the continuous bilinear form when we plug the exact solution $\widetilde{u}$ into the first argument and an arbitrary function from the discrete space V_h in the second argument,

$$\widetilde{a}_h^{(0)}(\widetilde{u}, \varphi_h) = \widetilde{a}(\widetilde{u}, \varphi_h) \qquad \forall \varphi_h \in V_h.$$

Owing to Lemma 2.14 this property is satisfied. Secondly, we know by Proposition 1.14 that the continuous bilinear form is skew-adjoint on $V_\star$,

$$\widetilde{a}(v_1, v_2) = -\widetilde{a}(v_2, v_1) \qquad \forall v_1, v_2 \in V_\star,$$

and consequently it holds,

$$\widetilde{a}(v, v) = 0 \qquad \forall v \in V_\star. \tag{3.9}$$

It is natural to demand this property also for the discrete bilinear form, but now on the discrete space V_h. Indeed, it turns out that (3.9) is crucial for the later analysis. Thus, we investigate if $\widetilde{a}_h^{(0)}$ satisfies

this property. Let therefore $v_h = [H_h, E_h]^T \in V_h$. Integrating by parts in the first term of $\widetilde{a}_h^{(0)}$ we deduce

$$\begin{aligned}\widetilde{a}_h^{(0)}(v_h, v_h) =& c_0(\nabla_h \times E_h, H_h)_{L^2(\Omega)^3} - c_0(\nabla_h \times H_h, E_h)_{L^2(\Omega)^3} \\ =& c_0 \sum_{K\in\mathcal{T}_h} \left[(\nabla \times E_K, H_K)_{L^2(K)^3} - (\nabla \times H_K, E_K)_{L^2(K)^3}\right] \\ =& c_0 \sum_{K\in\mathcal{T}_h} \left[(E_K, \nabla \times H_K)_{L^2(K)^3} - (\nabla \times H_K, E_K)_{L^2(K)^3}\right] \\ &+ c_0 \sum_{K\in\mathcal{T}_h} \sum_{F\in\mathcal{F}_K} (n_K \times E_K, H_K)_{L^2(F)^3} \\ =& c_0 \sum_{K\in\mathcal{T}_h} \sum_{F\in\mathcal{F}_K} (n_K \times E_K, H_K)_{L^2(F)^3}. \end{aligned} \tag{3.10}$$

Here we have dropped the index h in writing E_K and H_K instead of $E_{h,K} = (E_h)|_K$ and $H_{h,K} = (H_h)|_K$ to simplify the notation. We will stick to this notation whenever no confusion can arise. We clearly see that $\widetilde{a}_h^{(0)}(v_h, v_h) \neq 0$ and thus we have to change the ansatz. The easiest way to initiate the requested property is to substract the distracting term if we can ensure that consistency retains. From (3.10) this is not seen and thus we rewrite it. Recall that we defined for a mesh element K with interface F the element K_F as the neighboring mesh element w. r. t. F and n_F as the normal pointing from K to K_F, see Figure 2.2. Using this notation we can write

$$\begin{aligned}\sum_{K\in\mathcal{T}_h} \sum_{F\in\mathcal{F}_K} (n_K \times E_K, H_K)_{L^2(F)^3} =& \sum_{F\in\mathcal{F}_h^i} \left[(n_F \times E_K, H_K)_{L^2(F)^3} - (n_F \times E_{K_F}, H_{K_F})_{L^2(F)^3}\right] \\ &+ \sum_{F\in\mathcal{F}_h^b} (n \times E_h, H_h)_{L^2(F)^3}.\end{aligned}$$

The two summands in the first sum can further be rewritten as

$$(n_F \times E_K, H_K)_{L^2(F)^3} = \frac{1}{2}\left[(n_F \times E_K, H_K + H_{K_F})_{L^2(F)^3} + (n_F \times E_K, H_K - H_{K_F})_{L^2(F)^3}\right],$$

and

$$-(n_F \times E_{K_F}, H_{K_F})_{L^2(F)^3} = -\frac{1}{2}\left[(n_F \times E_{K_F}, H_K + H_{K_F})_{L^2(F)^3} + (n_F \times E_{K_F}, H_{K_F} - H_K)_{L^2(F)^3}\right].$$

Adding this two equations yields

$$\begin{aligned}(n_F \times E_K, H_K)_{L^2(F)^3} &- (n_F \times E_{K_F}, H_{K_F})_{L^2(F)^3} \\ &= -(n_F \times [\![E_h]\!]_F, \{\!\{H_h\}\!\}_F)_{L^2(F)^3} - (n_F \times \{\!\{E_h\}\!\}_F, [\![H_h]\!]_F)_{L^2(F)^3}.\end{aligned}$$

Using the vector identity $(n \times e) \cdot h = -(n \times h) \cdot e$ in the second term we finally obtain

$$\begin{aligned}\sum_{K\in\mathcal{T}_h} \sum_{F\in\mathcal{F}_K} (n_K \times E_K, H_K)_{L^2(F)^3} =& \sum_{F\in\mathcal{F}_h^i} \left[-(n_F \times [\![E_h]\!]_F, \{\!\{H_h\}\!\}_F)_{L^2(F)^3} + (n_F \times [\![H_h]\!]_F, \{\!\{E_h\}\!\}_F)_{L^2(F)^3}\right] \\ &+ \sum_{F\in\mathcal{F}_h^b} (n \times E_h, H_h)_{L^2(F)^3}.\end{aligned} \tag{3.11}$$

Recalling properties (3.6), (3.7) stemming from Assumption 3.2, we see that for the exact solution (3.11) vanishes,

$$\sum_{K\in\mathcal{T}_h} \sum_{F\in\mathcal{F}_K} (n_K \times \widetilde{\mathcal{E}}_K, \widetilde{\mathcal{H}}_K)_{L^2(F)^3} = 0.$$

Consequently, we have the following definition.

Definition 3.3 (Centered fluxes bilinear form). We define the discrete *centered fluxes bilinear form* $\widetilde{a}_h^{\mathrm{cf}} : V_{\star h} \times V_h \to \mathbb{R}$ as follows: For $v = [H, E]^T$ and $\varphi_h = [\phi_h, \psi_h]^T$,

$$\begin{aligned}\widetilde{a}_h^{\mathrm{cf}}(v, \varphi_h) :=& c_0(\nabla_h \times E, \phi_h)_{L^2(\Omega)^3} - c_0(\nabla_h \times H, \psi_h)_{L^2(\Omega)^3} \\ &+ c_0 \sum_{F \in \mathcal{F}_h^i} \left[(n_F \times [\![E]\!]_F, \{\!\{\phi_h\}\!\}_F)_{L^2(F)^3} - (n_F \times [\![H]\!]_F, \{\!\{\psi_h\}\!\}_F)_{L^2(F)^3} \right] \\ &+ c_0 \sum_{F \in \mathcal{F}_h^b} \left[-(n \times E, \phi_h)_{L^2(F)^3} \right]. \end{aligned} \tag{3.12}$$

We will see later where the name 'centered fluxes' stems from. Now, we can state the discretization of (3.4): We search for $\widetilde{u}_h = [\widetilde{\mathcal{H}}_h, \widetilde{\mathcal{E}}_h]^T \in C^1(0, T; V_h)$ such that

$$\widetilde{m}_h(\partial_t \widetilde{u}_h, \varphi_h) + \widetilde{a}_h^{\mathrm{cf}}(\widetilde{u}_h, \varphi_h) = (\widetilde{g}, \varphi_h)_{L^2(\Omega)^6} \qquad \forall \varphi_h \in V_h. \tag{3.13}$$

The next lemma collects properties of $\widetilde{a}_h^{\mathrm{cf}}$.

Lemma 3.4 *(Consistency and skew-adjointness). The discrete bilinear form $\widetilde{a}_h^{\mathrm{cf}}$ satisfies the following properties:*

i) Consistency*, i. e. for the exact solution $\widetilde{u} \in V_\star$ it holds*

$$\widetilde{a}_h^{\mathrm{cf}}(\widetilde{u}, \varphi_h) = \widetilde{a}(\widetilde{u}, \varphi_h) \qquad \forall \varphi_h \in V_h. \tag{3.14}$$

Indeed, this holds true for every $v \in V_\star$.

ii) Skew-adjointness *on V_h, i. e.*

$$\widetilde{a}_h^{\mathrm{cf}}(v_h, \widehat{v}_h) = -\widetilde{a}_h^{\mathrm{cf}}(\widehat{v}_h, v_h) \qquad \forall v_h, \widehat{v}_h \in V_h. \tag{3.15}$$

We see that despite having constructed $\widetilde{a}_h^{\mathrm{cf}}$ only to satisfy $\widetilde{a}_h^{\mathrm{cf}}(v_h, v_h) = 0$, we get a stronger property, namely the skew-adjointness property.

Proof: i) By construction.
ii) Let $v_h = [H_h, E_h]^T, \widehat{v}_h = [\widehat{H}_h, \widehat{E}_h]^T \in V_h$. We integrate by parts the curl terms in (3.12)

$$\begin{aligned}(\nabla_h \times E_h, \widehat{H}_h)_{L^2(\Omega)^3} &- (\nabla_h \times H_h, \widehat{E}_h)_{L^2(\Omega)^3} \\ &= (E_h, \nabla_h \times \widehat{H}_h)_{L^2(\Omega)^3} - (H_h, \nabla_h \times \widehat{E}_h)_{L^2(\Omega)^3} \\ &\quad + \sum_{K \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_K} \left[(n_K \times E_K, \widehat{H}_K)_{L^2(F)^3} - (n_K \times H_K, \widehat{E}_K)_{L^2(F)^3} \right]. \end{aligned}$$

Using (3.11) we can write the last sum as

$$\begin{aligned}&\sum_{F \in \mathcal{F}_h^i} \left[-(n_F \times [\![E_h]\!]_F, \{\!\{\widehat{H}_h\}\!\}_F)_{L^2(F)^3} + (n_F \times [\![\widehat{H}_h]\!]_F, \{\!\{E_h\}\!\}_F)_{L^2(F)^3} \right] \\ &+ \sum_{F \in \mathcal{F}_h^i} \left[(n_F \times [\![H_h]\!]_F, \{\!\{\widehat{E}_h\}\!\}_F)_{L^2(F)^3} - (n_F \times [\![\widehat{E}_h]\!]_F, \{\!\{H_h\}\!\}_F)_{L^2(F)^3} \right] \\ &+ \sum_{K \in \mathcal{F}_h^b} \left[(n \times E_h, \widehat{H}_h)_{L^2(F)^3} - (n \times H_h, \widehat{E}_h)_{L^2(F)^3} \right]. \end{aligned}$$

Employing this in (3.12) then yields

$$\begin{aligned}\widetilde{a}_h^{\mathrm{cf}}(v_h, \widehat{v}_h) =& c_0(E_h, \nabla_h \times \widehat{H}_h)_{L^2(\Omega)^3} - c_0(H_h, \nabla_h \times \widehat{E}_h)_{L^2(\Omega)^3} \\ &+ c_0 \sum_{F \in \mathcal{F}_h^i} \left[(n_F \times [\![\widehat{H}_h]\!]_F, \{\!\{E_h\}\!\}_F)_{L^2(F)^3} - (n_F \times [\![\widehat{E}_h]\!]_F, \{\!\{H_h\}\!\}_F)_{L^2(F)^3} \right] \\ &+ c_0 \sum_{K \in \mathcal{F}_h^b} \left[-(n \times H_h, \widehat{E}_h)_{L^2(F)^3} \right]. \\ =& -\widetilde{a}_h^{\mathrm{cf}}(\widehat{v}_h, v_h) \end{aligned} \tag{3.16}$$

□

Fluxes

Starting from (3.16) we have the following equivalent representation of $\widetilde{a}_h^{\mathrm{cf}}$ where we use the convention $n_F \times [\![\psi_h]\!]_F = -n \times \psi_h$ for boundary faces $F \in \mathcal{F}_h^b$.

Lemma 3.5 *(Flux form) The discrete bilinear form $\widetilde{a}_h^{\mathrm{cf}}$ can be equivalently written as: For $v = [H, E]^T \in V_{\star h}$ and $\varphi_h = [\phi_h, \psi_h]^T \in V_h$,*

$$\begin{aligned}\widetilde{a}_h^{\mathrm{cf}}(v, \varphi_h) =& c_0(E, \nabla_h \times \phi_h)_{L^2(\Omega)^3} - c_0(H, \nabla_h \times \psi_h)_{L^2(\Omega)^3} \\ &+ c_0 \sum_{F \in \mathcal{F}_h} \left[(\widetilde{f}_\phi^{\mathrm{cf}}(H, E), n_F \times [\![\phi_h]\!]_F)_{L^2(F)^3} - (\widetilde{f}_\psi^{\mathrm{cf}}(H, E), n_F \times [\![\psi_h]\!]_F)_{L^2(F)^3} \right], \end{aligned} \tag{3.17}$$

with centered fluxes $\widetilde{f}_\phi^{\mathrm{cf}}$, $\widetilde{f}_\psi^{\mathrm{cf}}$ *defined as*

$$\widetilde{f}_\phi^{\mathrm{cf}}(H, E) := \begin{cases} \{\!\{E\}\!\}_F, & F \in \mathcal{F}_h^i, \\ 0, & F \in \mathcal{F}_h^b, \end{cases} \qquad \widetilde{f}_\psi^{\mathrm{cf}}(H, E) := \begin{cases} \{\!\{H\}\!\}_F, & F \in \mathcal{F}_h^i, \\ H, & F \in \mathcal{F}_h^b. \end{cases}$$

Owing to the discontinuous ansatz the first two terms in (3.17) do not admit a transfer of information between the mesh elements. This task is accomplished by the flux functions. We see that in the current case we couple two neighboring elements by their meanvalue on the shared interface. This explains the name 'centered fluxes' for this scheme. The advantage of the flux notation is that we gain the freedom of choosing different fluxes in order to obtain an enhanced discretization. An indicator of the quality of the discretization is the residual of the discrete solution.

Local residuals

Recalling that the continuous problem (3.2) is set in the space $V_\star$ and that we are working with a non-comforming method with $V_h \notin V_\star$, we realize that we cannot define a global residual. However, if we consider the continuous problem on a single mesh-element K this incompatibility vanishes since the discrete solution is locally smooth, namely in $V_h(K) := \mathbb{P}_3^k(K)^3 \times \mathbb{P}_3^k(K)^3$. This motivates the introduction of local residuals.

Definition 3.6 (Local residuals). Let $\widetilde{u}_h = [\widetilde{\mathcal{H}}_h, \widetilde{\mathcal{E}}_h]^T \in V_h$ be the discrete solution obtained from (3.13). Furthermore, let $K \in \mathcal{T}_h$ be a mesh element. We define the *local residual* $\widetilde{r}_K = [\widetilde{r}_K^{\mathcal{H}}, \widetilde{r}_K^{\mathcal{E}}] \in V_h(K)$ on the mesh element K as

$$\widetilde{r}_K^{\mathcal{H}} := \partial_t \widetilde{\mathcal{H}}_{h,K} + c_0 \nabla \times \widetilde{\mathcal{E}}_{h,K}, \tag{3.18a}$$

$$\widetilde{r}_K^{\mathcal{E}} := \partial_t \widetilde{\mathcal{E}}_{h,K} - c_0 \nabla \times \widetilde{\mathcal{H}}_{h,K} + \widetilde{\mathcal{J}}_K. \tag{3.18b}$$

Using the characteristic function χ_K on the mesh element K, defined by

$$\chi_K(x) := \begin{cases} 1 & \text{if } x \in K, \\ 0 & \text{if } x \notin K, \end{cases}$$

we can connect the local residual to the continuous bilinear forms $\widetilde{m}$ and $\widetilde{a}$: For all $\varphi_h = [\phi_h, \psi_h]^T \in V_h$ it holds

$$\begin{aligned}(\widetilde{r}_K, \chi_K \varphi_h)_{L^2(K)^6} =& (\partial_t \widetilde{\mathcal{H}}_h, \chi_K \phi_h)_{L^2(K)^3} + c_0 (\nabla \times \widetilde{\mathcal{E}}_h, \chi_K \phi_h)_{L^2(K)^3} \\ &+ (\partial_t \widetilde{\mathcal{E}}_h, \chi_K \psi_h)_{L^2(K)^3} - c_0 (\nabla \times \widetilde{\mathcal{H}}_h, \chi_K \psi_h)_{L^2(K)^3} \\ &+ (\widetilde{\mathcal{J}}, \chi_K \psi_h)_{L^2(K)^3} \\ =& \widetilde{m}(\partial_t \widetilde{u}_h, \chi_K \varphi_h) + \widetilde{a}(\widetilde{u}_h, \chi_K \varphi_h) - (\widetilde{g}, \chi_K \varphi_h)_{L^2(\Omega)^6}. \end{aligned} \tag{3.19}$$

This representation enables us to draw a link between the local residual and the discrete problem (3.13) and reveal that the local residual is closely connected to the tangential jumps of the discrete solution.

Lemma 3.7 *(Local residuals and tangential jumps). For the local residual* $\widetilde{r}_K = [\widetilde{r}_K^{\mathcal{H}}, \widetilde{r}_K^{\mathcal{E}}] \in V_h(K)$ *there holds*

$$\begin{aligned}\widetilde{r}_K^{\mathcal{H}} &= -\frac{1}{2}c_0 \sum_{F\in\mathcal{F}_K^i} n_F \times [\![\widetilde{\mathcal{E}}_h]\!]_F + c_0 \sum_{F\in\mathcal{F}_K^b} n\times\widetilde{\mathcal{E}}_h,\\ \widetilde{r}_K^{\mathcal{E}} &= \frac{1}{2}c_0 \sum_{F\in\mathcal{F}_K^i} n_F \times [\![\widetilde{\mathcal{H}}_h]\!]_F,\end{aligned}$$

where $\mathcal{F}_K^i := \mathcal{F}_K \cap \mathcal{F}_h^i$ *and* $\mathcal{F}_K^b := \mathcal{F}_K \cap \mathcal{F}_h^b$.

Proof: Since the broken curl coincides with the usual curl on every mesh element K we have for all $\varphi_h = [\phi_h, \psi_h]^T \in V_h$,

$$\begin{aligned}\widetilde{a}_h^{\mathrm{cf}}(\widetilde{u}_h, \chi_K\varphi_h) =& c_0(\nabla\times\widetilde{\mathcal{E}}_h, \chi_K\phi_h)_{L^2(K)^3} - c_0(\nabla\times\widetilde{\mathcal{H}}_h, \chi_K\psi_h)_{L^2(K)^3}\\ &+ \frac{1}{2}c_0 \sum_{F\in\mathcal{F}_K^i}\left[(n_F\times[\![\widetilde{\mathcal{E}}_h]\!]_F, \phi_K)_{L^2(F)^3} - (n_F\times[\![\widetilde{\mathcal{H}}_h]\!]_F, \psi_K)_{L^2(F)^3}\right]\\ &+ c_0 \sum_{F\in\mathcal{F}_K^b}\left[-(n\times\widetilde{\mathcal{E}}_h, \phi_K)_{L^2(F)^3}\right].\end{aligned}$$

The first two terms are just $\widetilde{a}(\widetilde{u}_h, \chi_K\varphi_h)$. Now we can conlcude from (3.19), (3.13) and $\widetilde{m}_h = \widetilde{m}$ that for all $\varphi_h \in V_h$ it holds

$$\begin{aligned}(\widetilde{r}_K, \chi_K\varphi_h)_{L^2(K)^6} =& -\frac{1}{2}c_0 \sum_{F\in\mathcal{F}_K^i}\left[(n_F\times[\![\widetilde{\mathcal{E}}_h]\!]_F, \phi_K)_{L^2(F)^3} - (n_F\times[\![\widetilde{\mathcal{H}}_h]\!]_F, \psi_K)_{L^2(F)^3}\right]\\ &+ c_0 \sum_{F\in\mathcal{F}_K^b}(n\times\widetilde{\mathcal{E}}_h, \phi_K)_{L^2(F)^3}.\end{aligned}$$

The assertion follows by choosing consecutively $\varphi_h = [\phi_h, 0]T$ and $\varphi_h = [0, \psi_h]^T$. □

Due to Lemma 3.7 small tangential jumps imply that the residuals are small, too, and therefore indicate a good approximate solution. In the next section we construct a discrete bilinear form with this property.

The upwind bilinear form $\widetilde{a}_h^{\mathrm{upw}}$

Motivated by the considerations above we want to improve the discrete bilinear form $\widetilde{a}_h^{\mathrm{cf}}$ by penalizing tangential jumps over interfaces. This can be achieved be replacing the centered fluxes $\widetilde{f}_\phi^{\mathrm{cf}}$, $\widetilde{f}_\psi^{\mathrm{cf}}$ in (3.17) by the so called *upwind fluxes*,

$$\begin{aligned}\widetilde{f}_\phi^{\mathrm{upw}}(H,E) &:= \begin{cases}\{\!\{E\}\!\}_F + \frac{1}{2}n_F\times[\![H]\!]_F, & F\in\mathcal{F}_h^i,\\ 0, & F\in\mathcal{F}_h^b,\end{cases}\\ \widetilde{f}_\psi^{\mathrm{upw}}(H,E) &:= \begin{cases}\{\!\{H\}\!\}_F - \frac{1}{2}n_F\times[\![E]\!]_F, & F\in\mathcal{F}_h^i,\\ H + n\times E, & F\in\mathcal{F}_h^b.\end{cases}\end{aligned}$$

Note that the upwind fluxes are just the centered fluxes plus some stabilization term, $\widetilde{f}_\phi^{\mathrm{upw}} = \widetilde{f}_\phi^{\mathrm{cf}} + \widetilde{f}_\phi^{\mathrm{s}}$, $\widetilde{f}_\psi^{\mathrm{upw}} = \widetilde{f}_\psi^{\mathrm{cf}} + \widetilde{f}_\psi^{\mathrm{s}}$. In [1] it is pointend out that this viewpoint, i. e. considering upwind fluxes as stabilization of the centered fluxes, is beneficial. In this spirit we directly define the improved form as sum of the centered fluxes bilinearform and a stabiliziation bilinearform.

Definition 3.8 (Upwind fluxes bilinear form). We define the discrete *upwind bilinear form* $\widetilde{a}_h^{\mathrm{upw}} : V_{\star h} \times V_h \to \mathbb{R}$ by: For $v = [H, E]^T$ and $\varphi_h = [\phi_h, \psi_h]^T$,

$$\widetilde{a}_h^{\mathrm{upw}}(v, \varphi) := \widetilde{a}_h^{\mathrm{cf}}(v, \varphi) + \widetilde{s}_h(v, \varphi_h), \tag{3.20}$$

where the *stabiliziation bilinear form* $\widetilde{s}_h : V_{\star h} \times V_h \to \mathbb{R}$ is given as,

$$\begin{aligned}\widetilde{s}_h(v, \varphi_h) :=& c_0 \sum_{F \in \mathcal{F}_h^i} \left[\frac{1}{2} (n_F \times [\![H]\!]_F, n_F \times [\![\phi_h]\!]_F)_{L^2(F)^3} + \frac{1}{2} (n_F \times [\![E]\!]_F, n_F \times [\![\psi_h]\!]_F)_{L^2(F)^3} \right] \\ &+ c_0 \sum_{F \in \mathcal{F}_h^b} (n \times E, \psi_h)_{L^2(F)^3}.\end{aligned} \tag{3.21}$$

The next lemma uncovers the properties of $\widetilde{a}_h^{\mathrm{upw}}$.

Lemma 3.9 *(Consistency and dissipation). The upwind bilinear form $\widetilde{a}_h^{\mathrm{upw}}$ satisfies the following properties:*

i) Consistency*, i. e. for the exact solution $\widetilde{u} \in V_\star$ there holds*

$$\widetilde{a}_h^{\mathrm{upw}}(\widetilde{u}, \varphi_h) = \widetilde{a}(\widetilde{u}, \varphi_h) \qquad \forall \varphi_h \in V_h. \tag{3.22}$$

In fact, this property holds true for all $v \in V_\star$.

ii) Dissipation*, i. e. for the bilinear form $-\widetilde{a}_h^{\mathrm{upw}}$ it holds*

$$-\widetilde{a}_h^{\mathrm{upw}}(v_h, v_h) = -\widetilde{s}_h(v_h, v_h) \leq 0 \qquad \forall v_h \in V_h. \tag{3.23}$$

We see that consistency is kept while skew-adjointness is replaced by the dissipative property (3.23). Indeed, this property turns out to be crucial for the better convergence of the upwind fluxes discretization.

Proof: i) Consistency of $\widetilde{a}_h^{\mathrm{cf}}$ implies that for all $\varphi_h \in V_h$ there holds

$$\widetilde{a}_h^{\mathrm{upw}}(\widetilde{u}, \varphi_h) = \widetilde{a}_h^{\mathrm{cf}}(\widetilde{u}, \varphi_h) + \widetilde{s}_h(\widetilde{u}, \varphi_h) = \widetilde{a}(\widetilde{u}, \varphi_h) + \widetilde{s}_h(\widetilde{u}, \varphi_h).$$

Applying (3.6) and (3.7) to the stabilization form reveals

$$\widetilde{s}_h(\widetilde{u}, \varphi_h) = 0 \qquad \forall \varphi_h \in V_h,$$

and thus we infer (3.22). The same arguments are valid for all functions $v \in V_\star$.
ii) Let $v_h = [H_h, E_h]^T \in V_h$. Then, using the skew-adjointness of $\widetilde{a}_h^{\mathrm{cf}}$, we have

$$\begin{aligned}\widetilde{a}_h^{\mathrm{upw}}(v_h, v_h) =& \widetilde{s}_h(v_h, v_h) = \frac{1}{2} c_0 \sum_{F \in \mathcal{F}_h^i} \left[\| n_F \times [\![H_h]\!] \|^2_{L^2(F)^3} + \| n_F \times [\![E_h]\!]_F \|^2_{L^2(F)^3} \right] \\ &+ c_0 \sum_{F \in \mathcal{F}_h^b} \| n \times E_h \|^2_{L^2(F)^3} \geq 0.\end{aligned}$$

□

This ends the construction of the discrete bilinear forms. Next, we will show how this ideas can be carried over to the case of non-constant coefficients ε and μ.

3.2 Inhomogeneous Medium

In this section we consider inhomogeneous media, in particular we consider the case of composite media, i. e. the material coefficients ε and μ are piecewise constant. We make the assumption that the mesh is matched to the coefficients in the sense that on every mesh element the coefficients are constant

$$\varepsilon|_K,\, \mu|_K \equiv const \qquad \forall K \in \mathcal{T}_h.$$

Clearly, we cannot use the normalized Maxwell's equations (3.2) anymore but have to work with the standard equations

$$\partial_t \mathcal{H} + \mu^{-1} \nabla \times \mathcal{E} = 0 \qquad \text{in } (0,T)\times\Omega, \tag{3.24a}$$

$$\partial_t \mathcal{E} - \varepsilon^{-1} \nabla \times \mathcal{H} = -\varepsilon^{-1} \mathcal{J} \qquad \text{in } (0,T)\times\Omega. \tag{3.24b}$$

Analogously to the homogeneous case the starting point of the discretization is the equivalent formulation of (3.24) with bilinear forms.

Definition 3.10 (Continuous bilinear forms). We define the bilinear forms $m,\, a : \mathcal{D}(A) \times V \to \mathbb{R}$ as follows: For $v = [H,E]^T$ and $\varphi = [\phi,\psi]^T$,

$$\begin{aligned} m(v,\varphi) &:= (\mu H, \phi)_{L^2(\Omega)^3} + (\varepsilon E, \psi)_{L^2(\Omega)^3}, \\ a(v,\varphi) &:= (\nabla \times E, \phi)_{L^2(\Omega)^3} - (\nabla \times H, \psi)_{L^2(\Omega)^3}. \end{aligned}$$

Note that in this section we work with the V-inner product and consequently the bilinear forms stem from taking the V-inner product from (3.24) with test functions $\varphi \in V$. We emphasize that the bilinear-forms $\widetilde{a}$ and a coincide except for the (constant) factor c_0 and thus the calculations carried out for $\widetilde{a}_h^{\mathrm{cf}}$ can be used in the construction of the discretization of a.

Furthermore, we can equivalently state Maxwell's equations (3.24) with the bilinear forms as: We search for $u = [\mathcal{H},\mathcal{E}]^T \in C^1(0,T;V) \cap C(0,T;\mathcal{D}(A))$ such that

$$m(\partial_t u, \varphi) + a(u,\varphi) = (g,\varphi)_V \qquad \forall \varphi \in V, \tag{3.25}$$

where the source term is $g = [0, -\varepsilon\mathcal{J}]^T$.

3.2.1 Discrete Bilinear Forms

Let Assumption 3.2 hold for the exact solution u of (3.25). We can yet define the discrete bilinear form $m_h : V_{\star h} \times V_h \to \mathbb{R}$: For $v = [H,E]^T$ and $\varphi_h = [\phi_h, \psi_h]^T$,

$$m_h(v,\varphi_h) := (\mu H, \phi_h)_{L^2(\Omega)^3} + (\varepsilon E, \psi_h)_{L^2(\Omega)^3}. \tag{3.26}$$

Centered fluxes

As explained above we can use the calculations proven for the case of homogeneous media. In particular, we can use the flux form derived in Lemma 3.5 by just dropping the factor c_0. We use this form as starting point for the construction of the centered flux bilinear form: For $v = [H,E]^T \in V_{\star h}$ and $\varphi_h = [\phi_h,\psi_h]^T \in V_h$ we define

$$\begin{aligned} a_h^{\mathrm{cf}}(v,\varphi_h) :=& (E, \nabla_h \times \phi_h)_{L^2(\Omega)^3} - (H, \nabla_h \times \psi_h)_{L^2(\Omega)^3} \\ &+ \sum_{F \in \mathcal{F}_h} \left[(f_\phi^{\mathrm{cf}}(H,E), n_F \times [\![\phi_h]\!]_F)_{L^2(F)^3} - (f_\psi^{\mathrm{cf}}(H,E), n_F \times [\![\psi_h]\!]_F)_{L^2(F)^3} \right]. \end{aligned} \tag{3.27}$$

Recall that in the case of homogeneous media we chose the fluxes on the interface F as the average of the two-valued trace of the functions H and E stemming from the two neighboring mesh elements K and K_F composing this inferface. In the current case of composite media we want to reflect the fact that every mesh element can posses different material coefficients ε and μ by using weighted averages.

Definition 3.11 (Weighted averages). Let ω_K and ω_{K_F} be positive weights associated with the element K and K_F, respectively. Let v be a piecewise smooth function on $K \cup K_F$ which admits a trace on F. Then, we define the *weighted average* w. r. t. to $\omega = \{\omega_K, \omega_{K_F}\}$ as

$$\{\!\{v\}\!\}_F^{\omega} := \frac{\omega_K v_K + \omega_{K_F} v_{K_F}}{\omega_K + \omega_{K_F}}.$$

Further, we define the *adjoint average* as

$$\{\!\{v\}\!\}_F^{\overline{\omega}} := \frac{\omega_{K_F} v_K + \omega_K v_{K_F}}{\omega_K + \omega_{K_F}}.$$

Now let us make the following ansatz for the flux functions

$$f_\phi^{\mathrm{cf}}(H,E) := \begin{cases} \{\!\{E\}\!\}_F^{\alpha}, & F \in \mathcal{F}_h^i, \\ 0, & F \in \mathcal{F}_h^b, \end{cases} \qquad f_\psi^{\mathrm{cf}}(H,E) := \begin{cases} \{\!\{H\}\!\}_F^{\beta}, & F \in \mathcal{F}_h^i, \\ H, & F \in \mathcal{F}_h^b, \end{cases} \tag{3.28}$$

where the weights $\alpha = \{\alpha_K, \alpha_{K_F}\}$ and $\beta = \{\beta_K, \beta_{K_F}\}$ are to be determined. Recalling the construction of the discrete bilinear form $\widetilde{a}_h^{\mathrm{cf}}$ in the former section we check if our ansatz satisfies the two basic features, consistency and $a_h^{\mathrm{cf}}(v_h, v_h) = 0$ for all $v_h \in V_h$. We begin with consistency. Therefore, we integrate by parts the two curl terms in (3.27), which yields

$$\begin{aligned} a_h^{\mathrm{cf}}(v,\varphi_h) =& (\nabla_h \times E, \phi_h)_{L^2(\Omega)^3} - (\nabla_h \times H, \psi_h)_{L^2(\Omega)^3} \\ &+ \sum_{F \in \mathcal{F}_h^i} \left[(\{\!\{E\}\!\}_F^{\alpha}, n_F \times [\![\phi_h]\!]_F)_{L^2(F)^3} - (\{\!\{H\}\!\}_F^{\beta}, n_F \times [\![\psi_h]\!]_F)_{L^2(F)^3} \right] \\ &+ \sum_{F \in \mathcal{F}_h^b} (H, n \times \psi_h)_{L^2(\Omega)^3} \\ &+ \sum_{K \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_K} \left[(E_K, n_K \times \phi_K)_{L^2(F)^3} - (H_K, n_K \times \psi_K)_{L^2(F)^3} \right]. \end{aligned} \tag{3.29}$$

The last sum can be written as

$$\begin{aligned} &\sum_{F \in \mathcal{F}_h^i} \left[(E_K, n_F \times \phi_K)_{L^2(F)^3} - (E_{K_F}, n_F \times \phi_{K_F})_{L^2(F)^3} \right] \\ &- \sum_{F \in \mathcal{F}_h^i} \left[(H_K, n_F \times \psi_K)_{L^2(F)^3} - (H_{K_F}, n_F \times \psi_{K_F})_{L^2(F)^3} \right] \\ &+ \sum_{F \in \mathcal{F}_h^b} \left[(E, n \times \phi_h)_{L^2(F)^3} - (H, n \times \psi_h)_{L^2(F)^3} \right]. \end{aligned}$$

Inserting this equality into (3.29) yields

$$\begin{aligned} a_h^{\mathrm{cf}}(v,\varphi_h) =& (\nabla_h \times E, \phi_h)_{L^2(\Omega)^3} - (\nabla_h \times H, \psi_h)_{L^2(\Omega)^3} \\ &+ \sum_{F \in \mathcal{F}_h^i} \left[(\{\!\{E\}\!\}_F^{\alpha} - E_{K_F}, n_F \times \phi_{K_F})_{L^2(F)^3} - (\{\!\{E\}\!\}_F^{\alpha} - E_K, n_F \times \phi_K)_{L^2(F)^3} \right] \\ &- \sum_{F \in \mathcal{F}_h^i} \left[(\{\!\{H\}\!\}_F^{\beta} - H_{K_F}, n_F \times \psi_{K_F})_{L^2(F)^3} - (\{\!\{H\}\!\}_F^{\beta} - H_K, n_F \times \psi_K)_{L^2(F)^3} \right] \\ &+ \sum_{F \in \mathcal{F}_h^b} (E, n \times \phi_h)_{L^2(\Omega)^3} \end{aligned}$$

For the two summands in the first sum we have

$$\begin{aligned} &(\{\!\{E\}\!\}_F^{\alpha} - E_{K_F}, n_F \times \phi_{K_F})_{L^2(F)^3} - (\{\!\{E\}\!\}_F^{\alpha} - E_K, n_F \times \phi_K)_{L^2(F)^3} \\ &\quad = \left(\frac{\alpha_K (E_K - E_{K_F})}{\alpha_K + \alpha_{K_F}}, n_F \times \phi_{K_F} \right)_{L^2(F)^3} - \left(\frac{\alpha_{K_F} (E_{K_F} - E_K)}{\alpha_K + \alpha_{K_F}}, n_F \times \phi_K \right)_{L^2(F)^3}, \end{aligned}$$

and for the two summands in the second sum

$$(\{\!\{H\}\!\}_F^{\beta} - H_{K_F}, n_F \times \psi_{K_F})_{L^2(F)^3} - (\{\!\{H\}\!\}_F^{\beta} - H_K, n_F \times \psi_K)_{L^2(F)^3}$$
$$= \left(\frac{\beta_K (H_K - H_{K_F})}{\beta_K + \beta_{K_F}}, n_F \times \psi_{K_F}\right)_{L^2(F)^3} - \left(\frac{\beta_{K_F} (H_{K_F} - H_K)}{\beta_K + \beta_{K_F}}, n_F \times \psi_K\right)_{L^2(F)^3}.$$

Using the adjoint weights we can write this as

$$(\{\!\{E\}\!\}_F^{\alpha} - E_{K_F}, n_F \times \phi_{K_F})_{L^2(F)^3} - (\{\!\{E\}\!\}_F^{\alpha} - E_K, n_F \times \phi_K)_{L^2(F)^3} = -([\![E]\!]_F, n_F \times \{\!\{\phi_h\}\!\}_F^{\overline{\alpha}})_{L^2(F)^3},$$

and

$$(\{\!\{H\}\!\}_F^{\beta} - H_{K_F}, n_F \times \psi_{K_F})_{L^2(F)^3} - (\{\!\{H\}\!\}_F^{\beta} - H_K, n_F \times \psi_K)_{L^2(F)^3} = -([\![H]\!]_F, n_F \times \{\!\{\psi_h\}\!\}_F^{\overline{\beta}})_{L^2(F)^3}.$$

Alltogether, we have

$$\begin{aligned} a_h^{\mathrm{cf}}(v, \varphi_h) =& (\nabla_h \times E, \phi_h)_{L^2(\Omega)^3} - (\nabla_h \times H, \psi_h)_{L^2(\Omega)^3} \\ &+ \sum_{F \in \mathcal{F}_h^i} \left[(n_F \times [\![E]\!]_F, \{\!\{\phi_h\}\!\}_F^{\overline{\alpha}})_{L^2(F)^3} - (n_F \times [\![H]\!]_F, \{\!\{\psi_h\}\!\}_F^{\overline{\beta}})_{L^2(F)^3} \right] \\ &+ \sum_{F \in \mathcal{F}_h^b} \left[-(n \times E, \phi_h)_{L^2(F)^3} \right], \end{aligned} \tag{3.30}$$

whence consistency follows by (3.6) and (3.7) without any constraint on the weights α, β. Next, we analyze if a_h^{cf} satisfies $a_h^{\mathrm{cf}}(v_h, v_h) = 0$ for all $v_h \in V_h$. Integrating by parts only in the first term in (3.27) and performing the same computations as above yields

$$a_h^{\mathrm{cf}}(v_h, v_h) = \sum_{F \in \mathcal{F}_h^i} \left[(n_F \times [\![E_h]\!]_F, \{\!\{H_h\}\!\}_F^{\overline{\alpha}})_{L^2(F)^3} - (\{\!\{H_h\}\!\}_F^{\beta}, n_F \times [\![E_h]\!]_F)_{L^2(F)^3} \right].$$

In order to ensure that this sum vanishes we have to demand the conditions

$$\frac{\alpha_K}{\alpha_K + \alpha_{K_F}} = \frac{\beta_{K_F}}{\beta_K + \beta_{K_F}}, \qquad \frac{\alpha_{K_F}}{\alpha_K + \alpha_{K_F}} = \frac{\beta_K}{\beta_K + \beta_{K_F}} \tag{3.31}$$

on the weights. An adequate choice introduced in [8] satisfying the upper conditions, is to take the *local conductance* and the *local impedance*, respectively,

$$\alpha_K = \left(\frac{\varepsilon_K}{\mu_K}\right)^{1/2} = c_K \varepsilon_K, \qquad \beta_K = \left(\frac{\mu_K}{\varepsilon_K}\right)^{1/2} = c_K \mu_K, \tag{3.32}$$

where $c_K = (\varepsilon_K \mu_K)^{-1/2}$ is the local speed of light. With this choice we can construct the centered fluxes bilinear form.

Definition 3.12 (Centered fluxes bilinear form). We define the discrete *centered fluxes bilinear form* $a_h^{\mathrm{cf}} : V_{\star h} \times V_h \to \mathbb{R}$ as follows: For $v = [H, E]^T$ and $\varphi_h = [\phi_h, \psi_h]^T$,

$$\begin{aligned} a_h^{\mathrm{cf}}(v, \varphi_h) :=& (\nabla_h \times E, \phi_h)_{L^2(\Omega)^3} - (\nabla_h \times H, \psi_h)_{L^2(\Omega)^3} \\ &+ \sum_{F \in \mathcal{F}_h^i} \left[(n_F \times [\![E]\!]_F, \{\!\{\phi_h\}\!\}_F^{\overline{c\varepsilon}})_{L^2(F)^3} - (n_F \times [\![H]\!]_F, \{\!\{\psi_h\}\!\}_F^{\overline{c\mu}})_{L^2(F)^3} \right] \\ &+ \sum_{F \in \mathcal{F}_h^b} \left[-(n \times E, \phi_h)_{L^2(F)^3} \right]. \end{aligned} \tag{3.33}$$

Revisiting the construction of the centered fluxes bilinear form, we see that we can equivalently state it in the following form.

Lemma 3.13 *(Integration by parts form of a_h^{cf}). For $v = [H, E]^T \in V_{\star h}$ and $\varphi_h = [\phi_h, \psi_h]^T \in V_h$ there holds*

$$\begin{aligned} a_h^{\mathrm{cf}}(v, \varphi) =& (E, \nabla_h \times \phi_h)_{L^2(\Omega)^3} - (H, \nabla_h \times \psi_h)_{L^2(\Omega)^3} \\ &+ \sum_{F \in \mathcal{F}_h^i} \left[(\{\!\{E\}\!\}_F^{c\varepsilon}, n_F \times [\![\phi_h]\!]_F)_{L^2(F)^3} - (\{\!\{H\}\!\}_F^{c\mu}, n_F \times [\![\psi_h]\!]_F)_{L^2(F)^3} \right] \\ &+ \sum_{F \in \mathcal{F}_h^b} (H, n \times \psi_h)_{L^2(\Omega)^3}. \end{aligned} \tag{3.34}$$

Proof: This follows by construction, see equations (3.27), (3.28) with the choice of the weights (3.32). □

Let us collect the properties of a_h^{cf} . Again, despite of constructing a_h^{cf} only to satisfy $a_h^{\mathrm{cf}}(v_h, v_h) = 0$ for all $v_h \in V_h$ we have that a_h^{cf} is even skew-adjoint.

Lemma 3.14 *(Consistency and skew-adjointness). The centered fluxes bilinear form satisfies the following properties:*

i) Consistency*, i. e. for the exact solution $u \in V_\star$ it holds*

$$a_h^{\mathrm{cf}}(u, \varphi_h) = a(u, \varphi_h) \qquad \forall \varphi_h \in V_h. \tag{3.35}$$

In fact, this property holds true for all $v \in V_\star$.

ii) Skew-adjointness *on V_h, i. e.*

$$a_h^{\mathrm{cf}}(v_h, \widehat{v}_h) = -a_h^{\mathrm{cf}}(\widehat{v}_h, v_h) \qquad \forall v_h, \widehat{v}_h \in V_h. \tag{3.36}$$

Proof: i) By construction.
ii) We insert $v_h = [H_h, E_h]^T, \widehat{v}_h = [\widehat{H}_h, \widehat{E}_h]^T \in V_h$ in the partial integration form (3.34),

$$\begin{aligned} a_h^{\mathrm{cf}}(v_h, \widehat{v}_h) =& (E_h, \nabla_h \times \widehat{H}_h)_{L^2(\Omega)^3} - (H_h, \nabla_h \times \widehat{E}_h)_{L^2(\Omega)^3} \\ &+ \sum_{F \in \mathcal{F}_h^i} \left[(\{\!\{E_h\}\!\}_F^{c\varepsilon}, n_F \times [\![\widehat{H}_h]\!]_F)_{L^2(F)^3} - (\{\!\{H_h\}\!\}_F^{c\mu}, n_F \times [\![\widehat{E}_h]\!]_F)_{L^2(F)^3} \right] \\ &+ \sum_{F \in \mathcal{F}_h^b} (H_h, n \times \widehat{E}_h)_{L^2(\Omega)^3}. \end{aligned}$$

Using the condition for the weights (3.31), we see

$$\{\!\{E_h\}\!\}_F^{c\varepsilon} = \{\!\{E_h\}\!\}_F^{\overline{c\mu}}, \qquad \{\!\{H_h\}\!\}_F^{c\mu} = \{\!\{H_h\}\!\}_F^{\overline{c\varepsilon}},$$

and thus

$$\begin{aligned} a_h^{\mathrm{cf}}(v_h, \widehat{v}_h) =& -(\nabla_h \times \widehat{E}_h, H_h)_{L^2(\Omega)^3} + (\nabla_h \times \widehat{H}_h, E_h)_{L^2(\Omega)^3} \\ &- \sum_{F \in \mathcal{F}_h^i} \left[(n_F \times [\![\widehat{E}_h]\!]_F, \{\!\{H_h\}\!\}_F^{\overline{c\varepsilon}})_{L^2(F)^3} - (n_F \times [\![\widehat{H}_h]\!]_F, \{\!\{E_h\}\!\}_F^{\overline{c\mu}})_{L^2(F)^3} \right] \\ &- \sum_{F \in \mathcal{F}_h^b} \left[-(n \times \widehat{E}_h, H_h)_{L^2(\Omega)^3} \right] \\ =& -a_h^{\mathrm{cf}}(\widehat{v}_h, v_h). \end{aligned}$$

□

This lemma ensures that we have constructed a meaningful discrete bilinearform a_h^{cf} and thus we can discretize (3.25) as follows: We search for $u_h \in C^1(0, T; V_h)$ such that

$$m_h(\partial_t u_h, \varphi_h) + a_h^{\mathrm{cf}}(u_h, \varphi_h) = (g, \varphi_h)_V \qquad \forall \varphi_h \in V_h. \tag{3.37}$$

We proceed by computing the local residuals of the approximate solution of this discretization.

Local residuals

First we adapt the definition from the previous section.

Definition 3.15 (Local residual). Let $u_h = [\mathcal{H}_h, \mathcal{E}_h]^T \in V_h$ be the solution of the discrete problem (3.37) and let $K \in \mathcal{T}_h$ be a mesh element. We define the *local residual* $r_K = [r_K^{\mathcal{H}}, r_K^{\mathcal{E}}]^T \in V_h(K)$ on the element K as

$$r_K^{\mathcal{H}} := \partial_t \mathcal{H}_{h,K} + \mu_K^{-1} \nabla \times \mathcal{E}_{h,K}, \tag{3.38a}$$

$$r_K^{\mathcal{E}} := \partial_t \mathcal{E}_{h,K} - \varepsilon_K^{-1} \nabla \times \mathcal{H}_{h,K} + \varepsilon_K^{-1} \mathcal{J}_K. \tag{3.38b}$$

Now let us connect the local residual to the continuous problem (3.25). For $\varphi_h = [\phi_h, \psi_h]^T \in V_h$ there holds

$$\begin{aligned}(r_K, \chi_K \varphi_h)_{V(K)} =& (\mu \partial_t \mathcal{H}_h, \chi_K \phi_h)_{L^2(K)^3} + (\nabla \times \mathcal{E}_h, \chi_K \phi_h)_{L^2(K)^3} \\ &+ (\varepsilon \partial_t \mathcal{E}_h, \chi_K \psi_h)_{L^2(K)^3} - (\nabla \times \mathcal{H}_h, \chi_K \psi_h)_{L^2(K)^3} + (\mathcal{J}, \chi_K \psi_h)_{L^2(K)^3} \\ =& m(u_h, \chi_K \varphi_h) + a(u_h, \chi_K \varphi_h) - (g, \chi_K \varphi_h)_V. \end{aligned} \tag{3.39}$$

Again, we can prove the link between local residuals and tangential jumps.

Lemma 3.16 *(Local residual and tangential jumps). For the local residual* $r_K = [r_K^{\mathcal{H}}, r_K^{\mathcal{E}}]^T \in V_h(K)$ *there holds*

$$\begin{aligned} r_K^{\mathcal{H}} =& - c_K \sum_{F \in \mathcal{F}_K^i} \left[\frac{1}{2\{\!\{c\mu\}\!\}_F} n_F \times [\![\mathcal{E}_h]\!]_F \right] + c_K \sum_{F \in \mathcal{F}_K^b} \left[\frac{1}{c\mu} n \times \mathcal{E}_h \right], \\ r_K^{\mathcal{E}} =& c_K \sum_{F \in \mathcal{F}_K^i} \left[\frac{1}{2\{\!\{c\varepsilon\}\!\}_F} n_F \times [\![\mathcal{H}_h]\!]_F \right]. \end{aligned}$$

Proof: Let $\varphi_h = [\phi_h, \psi_h]^T \in V_h$. Inserting u_h and $\chi_K \varphi_h$ into a_h^{cf} yields

$$\begin{aligned} a_h^{\mathrm{cf}}(u_h, \chi_K \varphi_h) =& (\nabla \times \mathcal{E}_h, \chi_K \phi_h)_{L^2(K)^3} - (\nabla \times \mathcal{H}_h, \chi_K \psi_h)_{L^2(K)^3} \\ &+ \sum_{F \in \mathcal{F}_K^i} \left[\left(n_F \times [\![\mathcal{E}_h]\!]_F, \frac{c_{K_F} \varepsilon_{K_F}}{2\{\!\{c\varepsilon\}\!\}_F} \phi_K \right)_{L^2(F)^3} - \left(n_F \times [\![\mathcal{H}_h]\!]_F, \frac{c_{K_F} \mu_{K_F}}{2\{\!\{c\mu\}\!\}_F} \psi_K \right)_{L^2(F)^3} \right] \\ &+ \sum_{F \in \mathcal{F}_K^b} \left[-(n \times \mathcal{E}_h, \phi_h)_{L^2(F)^3} \right]. \end{aligned}$$

The first two terms are equal to $a(u_h, \chi_K \varphi_h)$. Using (3.39) and (3.37) we get

$$\begin{aligned} (r_K, \chi_K \varphi_h)_{V(K)} =& - \sum_{F \in \mathcal{F}_K^i} \left[\left(n_F \times [\![\mathcal{E}_h]\!]_F, \frac{c_{K_F} \varepsilon_{K_F}}{2\{\!\{c\varepsilon\}\!\}_F} \phi_K \right)_{L^2(F)^3} - \left(n_F \times [\![\mathcal{H}_h]\!]_F, \frac{c_{K_F} \mu_{K_F}}{2\{\!\{c\mu\}\!\}_F} \psi_K \right)_{L^2(F)^3} \right] \\ &- \sum_{F \in \mathcal{F}_K^b} \left[-(n \times \mathcal{E}_h, \phi_h)_{L^2(F)^3} \right]. \end{aligned} \tag{3.40}$$

Using the condition (3.31) on the weights yields

$$\frac{c_{K_F} \varepsilon_{K_F}}{2\{\!\{c\varepsilon\}\!\}_F} = \frac{c_K \mu_K}{2\{\!\{c\mu\}\!\}_F}, \qquad \frac{c_{K_F} \mu_{K_F}}{2\{\!\{c\mu\}\!\}_F} = \frac{c_K \varepsilon_K}{2\{\!\{c\varepsilon\}\!\}_F}.$$

Thus, we can write

$$\begin{aligned} &(r_K^{\mathcal{H}}, \mu_K \chi_K \phi_h)_{L^2(K)^3} + (r_K^{\mathcal{E}}, \varepsilon_K \chi_K \psi_h)_{L^2(K)^3} \\ &\quad = - c_K \sum_{F \in \mathcal{F}_h^i} \left(\frac{1}{2\{\!\{c\mu\}\!\}_F} n_F \times [\![\mathcal{E}_h]\!]_F, \mu_K \phi_K \right)_{L^2(F)^3} + c_K \sum_{F \in \mathcal{F}_K^b} \left(\frac{1}{\mu c} n \times \mathcal{E}_h, \mu_K \phi_K \right)_{L^2(F)^3} \\ &\qquad - c_K \sum_{F \in \mathcal{F}_h^i} \left(\frac{1}{2\{\!\{c\varepsilon\}\!\}_F} n_F \times [\![\mathcal{H}_h]\!]_F, \varepsilon_K \psi_K \right)_{L^2(F)^3} \end{aligned}$$

Choosing consecutively $\varphi_h = [\phi_h, 0]^T$ and $\varphi_h = [0, \psi_h]^T$ yields the assertion. □

Upwind fluxes

Following the idea of reducing the local residuals we modify the centered fluxes f_ϕ^{cf}, f_ψ^{cf} and obtain the following upwind fluxes

$$f_\phi^{\mathrm{upw}}(H,E) := \begin{cases} \{\!\{E\}\!\}_F^{c\varepsilon} + \frac{1}{2\{\!\{c\varepsilon\}\!\}_F} n_F \times [\![H]\!]_F, & F \in \mathcal{F}_h^i, \\ 0, & F \in \mathcal{F}_h^b, \end{cases}$$

$$f_\psi^{\mathrm{upw}}(H,E) := \begin{cases} \{\!\{H\}\!\}_F^{c\mu} - \frac{1}{2\{\!\{c\mu\}\!\}_F} n_F \times [\![E]\!]_F, & F \in \mathcal{F}_h^i, \\ H + \frac{1}{c\mu} n \times E, & F \in \mathcal{F}_h^b. \end{cases}$$

Definition 3.17 (Upwind bilinear form and stabilization bilinear form). We define the discrete *upwind bilinear form* $a_h^{\mathrm{upw}} : V_{\star h} \times V_h \to \mathbb{R}$ as follows: For $v = [H,E]^T$ and $\varphi_h = [\phi_h, \psi_h]^T$,

$$a_h^{\mathrm{upw}}(v, \varphi_h) := a_h^{\mathrm{cf}}(v, \varphi_h) + s_h(v, \varphi_h), \tag{3.41}$$

where the *stabilization bilinear form* $s_h : V_{\star h} \times V_h \to \mathbb{R}$ is given as

$$\begin{aligned} s_h(v, \varphi_h) := &\sum_{F \in \mathcal{F}_h^i} \left[\frac{1}{2\{\!\{c\varepsilon\}\!\}_F} (n_F \times [\![H]\!]_F, n_F \times [\![\phi_h]\!]_F)_{L^2(F)^3} \right] \\ &+ \sum_{F \in \mathcal{F}_h^i} \left[\frac{1}{2\{\!\{c\mu\}\!\}_F} (n_F \times [\![E]\!]_F, n_F \times [\![\psi_h]\!]_F)_{L^2(F)^3} \right] \\ &+ \sum_{F \in \mathcal{F}_h^b} \left[\frac{1}{c\mu} (n \times E, n \times \psi_h)_{L^2(F)^3} \right]. \end{aligned} \tag{3.42}$$

The discretization of (3.25) by the upwind bilinear form then reads: We search for $u_h \in C^1(0,T;V_h)$ such that

$$m_h(\partial_t u_h, \varphi_h) + a_h^{\mathrm{upw}}(u_h, \varphi_h) = (g, \varphi_h)_V \qquad \forall \varphi_h \in V_h. \tag{3.43}$$

For the further analysis the stabilization bilinear form is important. Clearly, it is symmetric on $V_h \times V_h$ and can be extended to a sysmmetric bilinear form on $V_{\star h} \times V_{\star h}$. Furthermore, we can easily see that it is positive semi-definite on $V_{\star h}$: For $v = [H,E]^T \in V_{\star h}$ it holds

$$\begin{aligned} s_h(v_h, v_h) = &\sum_{F \in \mathcal{F}_h^i} \left[\frac{1}{2\{\!\{c\varepsilon\}\!\}_F} \|n_F \times [\![H]\!]_F\|_{L^2(F)^3}^2 + \frac{1}{2\{\!\{c\mu\}\!\}_F} \|n_F \times [\![E]\!]_F\|_{L^2(F)^3}^2 \right] \\ &+ \sum_{F \in \mathcal{F}_h^b} \left[\frac{1}{c\mu} \|n \times E\|_{L^2(F)^3}^2 \right] \geq 0. \end{aligned}$$

Consequently, we can use s_h to build a seminorm.

Definition 3.18 (S-seminorm). We define the S-seminorm on $V_{\star h}$ by

$$|v|_S := (s_h(v,v))^{1/2} \qquad \forall v \in V_{\star h}. \tag{3.44}$$

Clearly, this is not a norm, e.g. from (3.6) and (3.7) we see that for all functions $v \in V_\star$ there holds

$$|v|_S = 0. \tag{3.45}$$

We finish this section by stating two lemmas immediately arising from the construction of a_h^{upw}.

Lemma 3.19 *(Integration by parts form of a_h^{upw}). For $v = [H,E]^T \in V_{\star h}$ and $\varphi_h = [\phi_h, \psi_h]^T \in V_h$ there holds*

$$\begin{aligned} a_h^{\mathrm{upw}}(v,\varphi_h) =& (E, \nabla_h \times \phi_h)_{L^2(\Omega)^3} - (H, \nabla_h \times \psi_h)_{L^2(\Omega)^3} \\ &+ \sum_{F\in\mathcal{F}_h^i} \left(\{\!\{E\}\!\}_F^{c\varepsilon} + \frac{1}{2\{\!\{c\varepsilon\}\!\}_F} n_F \times [\![H]\!]_F, n_F \times [\![\phi_h]\!]_F \right)_{L^2(F)^3} \\ &+ \sum_{F\in\mathcal{F}_h^i} \left(-\{\!\{H\}\!\}_F^{c\mu} + \frac{1}{2\{\!\{c\mu\}\!\}_F} n_F \times [\![E]\!]_F, n_F \times [\![\psi_h]\!]_F \right)_{L^2(F)^3} \\ &+ \sum_{F\in\mathcal{F}_h^b} \left(H + \frac{1}{c\mu} n \times E, n \times \psi_h \right)_{L^2(F)^3}. \end{aligned}$$

Proof: This follows immediately with the integration by parts form of a_h^{cf}, see Lemma 3.13. □

Lemma 3.20 *(Consistency and dissipation). The upwind bilinear form satisfies following properties:*

i) Consistency*, i. e. for the exact solution $u \in V_\star$ there holds*

$$a_h^{\mathrm{upw}}(u,\varphi_h) = a(u,\varphi_h) \qquad \forall \varphi_h \in V_h. \tag{3.46}$$

This property stays true for all $v \in V_\star$.

ii) Dissipation*, i. e. the bilinear form $-a_h^{\mathrm{upw}}$ satisfies*

$$-a_h^{\mathrm{upw}}(v_h,v_h) = -|v_h|_S^2 \le 0 \qquad \forall v_h \in V_h. \tag{3.47}$$

Proof: i) This is a consequence of the consistency of the centered flux bilinear form, see Lemma 3.14, and the fact that the S-seminorm vanishes for all functions in $V_\star$, see (3.45).
ii) This follows directly from the skew-adjointness of a_h^{cf}, see Lemma 3.14. □

The first step in proving the convergence of the constructed discretizations is to show the boundedness of the derived discrete bilinear forms.

3.3 Boundedness of Discrete Bilinearforms

We begin by an additional assumption on the mesh sequence.

Assumption 3.21 *(Quasi-uniform mesh sequence). We assume that the mesh sequence $\mathcal{T}_\mathcal{H}$ is quasi-uniform, meaning that there is a constant C_{qu} such that for all $h \in \mathcal{H}$ there holds*

$$\max_{K\in\mathcal{T}_h} h_K \le C_{\mathrm{qu}} \min_{K\in\mathcal{T}_h} h_K.$$

Furthermore, we introduce some abbreviations. We set ε_∞, μ_∞ and c_∞ to be the maximal values of the coefficients ε, μ and c, i. e.

$$\varepsilon_\infty := \max_{K\in\mathcal{T}_h} \varepsilon_K, \quad \mu_\infty := \max_{K\in\mathcal{T}_h} \mu_K, \quad c_\infty := \max_{K\in\mathcal{T}_h} c_K.$$

For a weight $\omega = \{\omega_K, \omega_{K_F}\}$ we set

$$\{\omega\}_F := \omega_K + \omega_{K_F} = 2\{\!\{\omega\}\!\}_F.$$

In addition, for a function $v = [H,E]^T$ we use the convention

$$\nabla \times v = \begin{bmatrix} \nabla \times H \\ \nabla \times E \end{bmatrix},$$

and analogously for $\nabla_h \times v$, $\{\!\{v\}\!\}_F$ and $[\![v]\!]_F$. Our first result is the boundedness of the centered fluxes bilinear form.

Theorem 3.22 *(Boundedness of a_h^{cf}). For the centered fluxes bilinear form we have for all $v \in V_{\star h}$ and for all $\varphi_h \in V_h$,*

$$|a_h^{\mathrm{cf}}(v,\varphi_h)| \le \left(c_\infty \|\nabla_h \times v\|_V + C_{\mathrm{bnd}} c_\infty^{1/2} h^{-1/2} |v|_S\right) \|\varphi_h\|_V, \tag{3.48}$$

with $C_{\mathrm{bnd}} = (\sqrt{2}N_\partial^{1/2}+1)C_{\mathrm{tr}}C_{\mathrm{qu}}^{-1/2}$ independet of h.

Proof: Let $v = [H,E]^T \in V_{\star h}$ and $\varphi_h = [\phi_h,\psi_h]^T \in V_h$. The proof proceeds in three steps. First, we bound the two curl terms appearing in $a_h^{\mathrm{cf}}(v,\varphi_h)$. We have

$$\begin{aligned}
&(\nabla_h \times E, \phi_h)_{L^2(\Omega)^3} - (\nabla_h \times H, \psi_h)_{L^2(\Omega)^3} \\
&\qquad = \sum_{K\in\mathcal{T}_h} \left(\begin{bmatrix} \nabla\times E \\ \nabla \times H \end{bmatrix}, \begin{bmatrix} \phi_h \\ -\psi_h \end{bmatrix} \right)_{L^2(K)^6} \\
&\qquad = \sum_{K\in\mathcal{T}_h} \varepsilon_K^{-1/2}\mu_K^{-1/2} \left(\begin{bmatrix} \varepsilon_K^{1/2}\nabla\times E \\ \mu_K^{1/2}\nabla \times H \end{bmatrix}, \begin{bmatrix} \mu_K^{1/2}\phi_h \\ -\varepsilon_K^{1/2}\psi_h \end{bmatrix} \right)_{L^2(K)^6} \\
&\qquad \le \sum_{K\in\mathcal{T}_h} c_K \left\| \begin{bmatrix} \varepsilon_K^{1/2}\nabla\times E \\ \mu_K^{1/2}\nabla \times H \end{bmatrix} \right\|_{L^2(K)^6} \left\| \begin{bmatrix} \mu_K^{1/2}\phi_h \\ -\varepsilon_K^{1/2}\psi_h \end{bmatrix} \right\|_{L^2(K)^6} \\
&\qquad = \sum_{K\in\mathcal{T}_h} c_K \|\nabla\times v\|_{V(K)} \|\varphi_h\|_{V(K)},
\end{aligned}$$

where the inequality is obtained by the Cauchy-Schwarz inequality. Applying the Cauchy-Schwarz inequality for sequences, see (A.2), it follows

$$\begin{aligned}
\sum_{K\in\mathcal{T}_h} c_K \|\nabla\times v\|_{V(K)} \|\varphi_h\|_{V(K)} \le & c_\infty \left(\sum_{K\in\mathcal{T}_h} \|\nabla\times v\|_{V(K)}^2 \right)^{1/2} \left(\sum_{K\in\mathcal{T}_h} \|\varphi_h\|_{V(K)}^2 \right)^{1/2} \\
= & c_\infty \|\nabla_h \times v\|_V \|\varphi_h\|_V.
\end{aligned}$$

Next we bound the sum over the interfaces in $a_h^{\mathrm{cf}}(v,\varphi_h)$, i. e. the terms

$$\sum_{F\in\mathcal{F}_h^i} \left[(n_F \times [\![E]\!]_F, \{\!\{\phi_h\}\!\}_F^{\overline{c\varepsilon}})_{L^2(F)^3} - (n_F \times [\![H]\!]_F, \{\!\{\psi_h\}\!\}_F^{\overline{c\mu}})_{L^2(F)^3} \right].$$

Recall that owing to condition (3.31) on the weights it holds

$$\{\!\{\phi_h\}\!\}_F^{\overline{c\varepsilon}} = \{\!\{\phi_h\}\!\}_F^{c\mu} = \frac{1}{\{c\mu\}_F}(c_K\mu_K\phi_K + c_{K_F}\mu_{K_F}\phi_{K_F}),$$

and

$$\{\!\{\psi_h\}\!\}_F^{\overline{c\mu}} = \{\!\{\psi_h\}\!\}_F^{c\varepsilon} = \frac{1}{\{c\varepsilon\}_F}(c_K\varepsilon_K\psi_K + c_{K_F}\varepsilon_{K_F}\psi_{K_F}).$$

Thus, a single summand can be written as

$$\begin{aligned}
&(n_F \times [\![E]\!]_F, \{\!\{\phi_h\}\!\}_F^{c\mu})_{L^2(F)^3} - (n_F \times [\![H]\!]_F, \{\!\{\psi_h\}\!\}_F^{c\varepsilon})_{L^2(F)^3} \\
&= \left(\begin{bmatrix} \frac{1}{\{c\mu\}_F^{1/2}} n_F \times [\![E]\!]_F \\ \frac{1}{\{c\varepsilon\}_F^{1/2}} n_F \times [\![H]\!]_F \end{bmatrix}, \begin{bmatrix} \frac{1}{\{c\mu\}_F^{1/2}}(c_K\mu_K\phi_K + c_{K_F}\mu_{K_F}\phi_{K_F}) \\ \frac{1}{\{c\varepsilon\}_F^{1/2}}(c_K\varepsilon_K\psi_K + c_{K_F}\varepsilon_{K_F}\psi_{K_F}) \end{bmatrix} \right)_{L^2(F)^6}.
\end{aligned}$$

Let us denote this summand with $\mathcal{S}_F$. The Cauchy-Schwarz inequality reveals

$$\sum_{F\in\mathcal{F}_h^i}\mathcal{S}_F \leq \left(\sum_{F\in\mathcal{F}_h^i}\left\|\begin{bmatrix}\frac{1}{\{c\mu\}_F^{1/2}}n_F\times[\![E]\!]_F\\ \frac{1}{\{c\varepsilon\}_F^{1/2}}n_F\times[\![H]\!]_F\end{bmatrix}\right\|^2_{L^2(F)^6}\right)^{1/2}\times$$
$$\left(\sum_{F\in\mathcal{F}_h^i}\left\|\begin{bmatrix}\frac{1}{\{c\mu\}_F^{1/2}}(c_K\mu_K\phi_K+c_{K_F}\mu_{K_F}\phi_{K_F})\\ \frac{1}{\{c\varepsilon\}_F^{1/2}}(c_K\varepsilon_K\psi_K+c_{K_F}\varepsilon_{K_F}\psi_{K_F})\end{bmatrix}\right\|^2_{L^2(F)^6}\right)^{1/2}.$$

The first factor on the RHS is equal to

$$\left(\sum_{F\in\mathcal{F}_h^i}\left[\frac{1}{\{c\mu\}_F}\|n_F\times[\![E]\!]_F\|^2_{L^2(F)^3}+\frac{1}{\{c\epsilon\}_F}\|n_F\times[\![H]\!]_F\|^2_{L^2(F)^3}\right]\right)^{1/2},$$

which clearly can be bounded by $|v|_S$. For the second factor we first notice that

$$\frac{c_K\mu_K}{\{c\mu\}_F^{1/2}}\leq(c_K\mu_K)^{1/2},\qquad \frac{c_{K_F}\mu_{K_F}}{\{c\mu\}_F^{1/2}}\leq(c_{K_F}\mu_{K_F})^{1/2},$$
$$\frac{c_K\varepsilon_K}{\{c\varepsilon\}_F^{1/2}}\leq(c_K\varepsilon_K)^{1/2},\qquad \frac{c_{K_F}\varepsilon_{K_F}}{\{c\varepsilon\}_F^{1/2}}\leq(c_{K_F}\varepsilon_{K_F})^{1/2}.$$

Then, using the triangle inequality and Young's inequality, we infer that the second factor can be estimated by

$$\left(\sum_{F\in\mathcal{F}_h^i}\left\|\begin{bmatrix}\frac{1}{\{c\mu\}_F^{1/2}}(c_K\mu_K\phi_K+c_{K_F}\mu_{K_F}\phi_{K_F})\\ \frac{1}{\{c\varepsilon\}_F^{1/2}}(c_K\varepsilon_K\psi_K+c_{K_F}\varepsilon_{K_F}\psi_{K_F})\end{bmatrix}\right\|^2_{L^2(F)^6}\right)^{1/2}$$
$$\leq\left(\sum_{F\in\mathcal{F}_h^i}\left[2c_K\left\|\begin{bmatrix}\mu_K^{1/2}\phi_K\\ \varepsilon_K^{1/2}\psi_K\end{bmatrix}\right\|^2_{L^2(F)^6}+2c_{K_F}\left\|\begin{bmatrix}\mu_{K_F}^{1/2}\phi_{K_F}\\ \varepsilon_{K_F}^{1/2}\psi_{K_F}\end{bmatrix}\right\|^2_{L^2(F)^6}\right]\right)^{1/2}. \tag{3.49}$$

Since μ_K and ε_K are constant on the mesh element K the discrete trace inequality (2.14) retains the same for the V-norm. So, we continue by applying it to (3.49), which yields the following upper bound

$$\sqrt{2}C_{\mathrm{tr}}c_\infty^{1/2}\left(\sum_{F\in\mathcal{F}_h^i}\left[h_K^{-1}\left\|\begin{bmatrix}\mu_K^{1/2}\phi_K\\ \varepsilon_K^{1/2}\psi_K\end{bmatrix}\right\|^2_{L^2(K)^6}+h_{K_F}^{-1}\left\|\begin{bmatrix}\mu_{K_F}^{1/2}\phi_{K_F}\\ \varepsilon_{K_F}^{1/2}\psi_{K_F}\end{bmatrix}\right\|^2_{L^2(K)^6}\right]\right)^{1/2}$$
$$=\sqrt{2}C_{\mathrm{tr}}c_\infty^{1/2}\left(\sum_{F\in\mathcal{F}_h^i}\left[h_K^{-1}\|\varphi_K\|^2_{V(K)}+h_{K_F}^{-1}\|\varphi_{K_F}\|^2_{V(K_F)}\right]\right)^{1/2}. \tag{3.50}$$

By Assumption 3.21 we infer that there holds for all $K\in\mathcal{T}_h$,

$$h_K^{-1}\leq C_{\mathrm{qu}}^{-1}h^{-1}.$$

Thus, we can further estimate (3.50) by

$$\sqrt{2}C_{\mathrm{tr}}C_{\mathrm{qu}}^{-1/2}c_\infty^{1/2}h^{-1/2}\left(\sum_{F\in\mathcal{F}_h^i}\left[\|\varphi_K\|^2_{V(K)}+\|\varphi_{K_F}\|^2_{V(K_F)}\right]\right)^{1/2}$$
$$\leq\sqrt{2}C_{\mathrm{tr}}C_{\mathrm{qu}}^{-1/2}c_\infty^{1/2}h^{-1/2}\left(\sum_{K\in\mathcal{T}_h}\mathrm{card}(\mathcal{F}_K)\|\varphi_K\|^2_{V(K)}\right)^{1/2}. \tag{3.51}$$

Recall that we have defined N_∂ as the maximal number of faces composing a mesh element and that N_∂ is uniformly bounded w. r. t. the meshsize h, see Lemma 2.16. Thus, we have the following upper bound for (3.51)

$$\sqrt{2}C_{\mathrm{tr}}C_{\mathrm{qu}}^{-1/2}N_\partial^{1/2}c_\infty^{1/2}h^{-1/2}\left(\sum_{K\in\mathcal{T}_h}\|\varphi_K\|_{V(K)}^2\right)^{1/2} = \sqrt{2}C_{\mathrm{tr}}C_{\mathrm{qu}}^{-1/2}N_\partial^{1/2}c_\infty^{1/2}h^{-1/2}\|\varphi_h\|_V.$$

Alltogether, we have shown the following bound for the sum over the interfaces

$$\sum_{F\in\mathcal{F}_h^i}\left[(n_F\times[\![E]\!]_F,\{\!\{\phi_h\}\!\}_F^{\overline{c\varepsilon}})_{L^2(F)^3}-(n_F\times[\![H]\!]_F,\{\!\{\psi_h\}\!\}_F^{\overline{c\mu}})_{L^2(F)^3}\right]\le\widetilde{C}h^{-1/2}|v|_S\|\varphi_h\|_V, \tag{3.52}$$

with $\widetilde{C}=\sqrt{2}C_{\mathrm{tr}}C_{\mathrm{qu}}^{-1/2}c_\infty^{1/2}N_\partial^{1/2}$.
Finally, we bound the last term in $a_h^{\mathrm{cf}}(v,\varphi_h)$, i. e. the sum over the boundary faces. With the same arguments as for the interfaces we deduce

$$\begin{aligned}
\sum_{F\in\mathcal{F}_h^b}(n\times E,\phi_h)_{L^2(F)^3} \le& \sum_{F\in\mathcal{F}_h^b}\left[\left(\frac{1}{(c\mu)^{1/2}}\|n\times E\|_{L^2(F)^3}\right)\left(c^{1/2}\|\mu^{1/2}\phi_h\|_{L^2(F)^3}\right)\right]\\
\le&\left(\sum_{F\in\mathcal{F}_h^b}\frac{1}{c\mu}\|n\times E\|_{L^2(F)^3}^2\right)^{1/2}\left(\sum_{F\in\mathcal{F}_h^b}c\|\mu^{1/2}\phi_h\|_{L^2(F)^3}^2\right)^{1/2}\\
\le&|v|_S\left(\sum_{F\in\mathcal{F}_h^b}c\|\mu^{1/2}\phi_h\|_{L^2(F)^3}^2\right)^{1/2}\\
\le&C_{\mathrm{tr}}C_{\mathrm{qu}}^{-1/2}c_\infty^{1/2}h^{-1/2}|v|_S\left(\sum_{K\in\mathcal{T}_h}\|\mu^{1/2}\phi_h\|_{L^2(K)^3}^2\right)^{1/2}\\
\le&C_{\mathrm{tr}}C_{\mathrm{qu}}^{-1/2}c_\infty^{1/2}h^{-1/2}|v|_S\|\varphi_h\|_V,
\end{aligned}\tag{3.53}$$

and the proof is finisched. □

Next we show a similar result for the stabilization bilinear form.

Theorem 3.23 *(Boundedness of s_h). Let $v\in V_{\star h}$ and $\varphi_h\in V_h$. Then, for the stabilization bilinear form there holds*

$$|s_h(v,\varphi_h)|\le C_{\mathrm{bnd}}c_\infty^{1/2}h^{-1/2}|v|_S\|\varphi_h\|_V, \tag{3.54}$$

where the constant is $C_{\mathrm{bnd}}=(\sqrt{2}N_\partial^{1/2}+1)C_{\mathrm{tr}}C_{\mathrm{qu}}^{-1/2}$.

Proof: It holds

$$\begin{aligned}
s_h(v,\varphi_h) =& \sum_{F\in\mathcal{F}_h^i}\left(\begin{bmatrix}\frac{1}{\{c\mu\}_F^{1/2}}n_F\times[\![E]\!]_F\\ \frac{1}{\{c\varepsilon\}_F^{1/2}}n_F\times[\![H]\!]_F\end{bmatrix},\begin{bmatrix}\frac{1}{\{c\mu\}_F^{1/2}}n_F\times[\![\psi_h]\!]_F\\ \frac{1}{\{c\varepsilon\}_F^{1/2}}n_F\times[\![\phi_h]\!]_F\end{bmatrix}\right)_{L^2(F)^6}\\
&+\sum_{F\in\mathcal{F}_h^b}\left(\frac{1}{(c\mu)^{1/2}}n\times E,\frac{1}{(c\mu)^{1/2}}n\times\psi_h\right)_{L^2(F)^3}\\
\le&\left(\sum_{F\in\mathcal{F}_h^i}\left\|\begin{bmatrix}\frac{1}{\{c\mu\}_F^{1/2}}n_F\times[\![E]\!]_F\\ \frac{1}{\{c\varepsilon\}_F^{1/2}}n_F\times[\![H]\!]_F\end{bmatrix}\right\|_{L^2(F)^6}^2\right)^{1/2}\left(\sum_{F\in\mathcal{F}_h^i}\left\|\begin{bmatrix}\frac{1}{\{c\mu\}_F^{1/2}}n_F\times[\![\psi_h]\!]_F\\ \frac{1}{\{c\varepsilon\}_F^{1/2}}n_F\times[\![\phi_h]\!]_F\end{bmatrix}\right\|_{L^2(F)^6}^2\right)^{1/2}\\
&+\left(\sum_{F\in\mathcal{F}_h^b}\left\|\frac{1}{(c\mu)^{1/2}}n\times E\right\|_{L^2(F)^3}^2\right)^{1/2}\left(\sum_{F\in\mathcal{F}_h^b}\left\|\frac{1}{\{c\mu\}_F^{1/2}}n_F\times\psi_h\right\|_{L^2(F)^3}^2\right)^{1/2},
\end{aligned}$$

where we used the Cauchy-Schwarz inequality. We see that the respective first factors can be estimated by $|v|_S$ and further using $|n_F| = 1$ we get

$$s_h(v,\varphi_h) \leq |v|_S \left(\sum_{F\in\mathcal{F}_h^i} \left\| \begin{bmatrix} \frac{1}{\{c\mu\}_F^{1/2}}[\![\psi_h]\!]_F \\ \frac{1}{\{c\varepsilon\}_F^{1/2}}[\![\phi_h]\!]_F \end{bmatrix} \right\|^2_{L^2(F)^6} \right)^{1/2} + |v|_S \left(\sum_{F\in\mathcal{F}_h^b} \left\| \frac{1}{\{c\mu\}_F^{1/2}}\psi_h \right\|^2_{L^2(F)^3} \right)^{1/2}. \quad (3.55)$$

Clearly, there holds

$$\frac{1}{\{c\mu\}_F} \leq \frac{1}{c_K\mu_K} = c_K\varepsilon_K, \qquad \frac{1}{\{c\mu\}_F} \leq \frac{1}{c_{K_F}\mu_{K_F}} = c_{K_F}\varepsilon_{K_F},$$
$$\frac{1}{\{c\varepsilon\}_F} \leq \frac{1}{c_K\varepsilon_K} = c_K\mu_K, \qquad \frac{1}{\{c\varepsilon\}_F} \leq \frac{1}{c_{K_F}\varepsilon_{K_F}} = c_{K_F}\mu_{K_F},$$

and $(c\mu)^{-1} = c\varepsilon$. Using this property in (3.55) together with Young's inequality yields

$$s_h(v,\varphi_h) \leq |v|_S \left(\sum_{F\in\mathcal{F}_h^i} \left[2c_K \left\| \begin{bmatrix} \varepsilon_K^{1/2}\psi_K \\ \mu_K^{1/2}\phi_K \end{bmatrix} \right\|^2_{L^2(F)^6} + 2c_{K_F} \left\| \begin{bmatrix} \varepsilon_{K_F}^{1/2}\psi_{K_F} \\ \mu_{K_F}^{1/2}\phi_{K_F} \end{bmatrix} \right\|^2_{L^2(F)^6} \right] \right)^{1/2}$$
$$+ |v|_S \left(\sum_{F\in\mathcal{F}_h^b} c \left\| \varepsilon^{1/2}\psi_h \right\|^2_{L^2(F)^3} \right)^{1/2}.$$

Using the same arguments as in the deduction from (3.49) to (3.53) in the previous proof we infer

$$s_h(v,\varphi_h) \leq (\sqrt{2}N_\partial^{1/2} + 1) C_{\mathrm{tr}} C_{\mathrm{qu}}^{-1/2} c_\infty^{1/2} h^{-1/2} |v|_S \|\varphi_h\|_V$$

□

We can easily draw the following corollary.

Corollary 3.24 *(Bound for S-seminorm). For all $v \in V_{\star h}$ there holds the following estimate*

$$|v|_S \leq C_{\mathrm{S}} h^{-1/2} \|v\|_V, \quad (3.56)$$

with $C_{\mathrm{S}} = C_{\mathrm{bnd}} c_\infty^{1/2}$.

Proof: Let $v \in V_{\star h}$ and let $v_\star \in V_\star$, $v_h \in V_h$ such that $v = v_\star + v_h$. In (3.45) we have shown that it holds $|v_\star|_S = 0$. Furthermore, using Theorem 3.23, we infer

$$|v_h|_S^2 = s_h(v_h, v_h) \leq C_{\mathrm{bnd}} c_\infty h^{-1/2} |v_h|_S \|v_h\|_V,$$

whence by dividing through $|v_h|_S$ we see

$$|v_h|_S \leq C_{\mathrm{bnd}} c_\infty h^{-1/2} \|v_h\|_V.$$

The assertion then follows by the triangle inequality. □

Theorems 3.22 and 3.23 immediately ensure the boundedness of the upwind bilinear form.

Theorem 3.25 *(Boundedness of a_h^{upw}). For all $v \in V_{\star h}$ and for all $\varphi_h \in V_h$ there holds*

$$|a_h^{\mathrm{upw}}(v,\varphi_h)| \leq (c_\infty \|\nabla_h \times v\|_V + 2C_{\mathrm{bnd}} c_\infty^{1/2} h^{-1/2} |v|_S) \|\varphi_h\|_V. \quad (3.57)$$

This concludes the section on the boundedness of the discrete bilinear forms. The next step is to derive an operator based approach.

3.4 Discrete Operators

We begin this section by recalling that we have formulated Maxwell's equations as the abstract evolution problem (see (1.35)): Search for $u \in C^1(0,T;\in V_\star) \cap C(0,T;\mathcal{D}(A))$ such that $u(0) = u_0$ and

$$\partial_t u + Au = g.$$

For the convergence analysis it is beneficial to write the centered fluxes and the upwind fluxes discretizations (3.37), (3.43) also in an operator based notation.

Definition 3.26 (Discrete operators). We define the operators A_h^{cf}, A_h^{upw}, $S_h : V_{\star h} \to V_h$ by

$$\begin{aligned}(A_h^{\mathrm{cf}} v, \varphi_h)_V &:= a_h^{\mathrm{cf}}(v,\varphi_h) \qquad \forall \varphi_h \in V_h,\\ (A_h^{\mathrm{upw}} v, \varphi_h)_V &:= a_h^{\mathrm{upw}}(v,\varphi_h) \qquad \forall \varphi_h \in V_h,\\ (S_h v, \varphi_h)_V &:= s_h(v,\varphi_h) \qquad \forall \varphi_h \in V_h.\end{aligned}$$

Obviously, there holds $A_h^{\mathrm{upw}} = A_h^{\mathrm{cf}} + S_h$. Furthermore, we introduce the projection onto V_h w. r. t. to the V-inner product.

Definition 3.27 (V-projection). We define the *V-projection onto V_h* as $\pi_h^V : V \to V_h$ such that

$$(\pi_h^V v, \varphi_h)_V = (v,\varphi_h)_V \qquad \forall \varphi_h \in V_h. \tag{3.58}$$

We work throughout the following sections with the V-inner product and thus ommit the index V and always assume that π_h denotes the V-projection π_h^V. Note that we have for all $v \in V$

$$\|\pi_h v\|_V = \sup_{\substack{\varphi_h \in V_h \\ \|\varphi_h\|_V = 1}} (\pi_h v, \varphi_h)_V = \sup_{\substack{\varphi_h \in V_h \\ \|\varphi_h\|_V = 1}} (v,\varphi_h)_V \le \sup_{\substack{\varphi_h \in V_h \\ \|\varphi_h\|_V = 1}} \|v\|_V \|\varphi_h\|_V = \|v\|_V. \tag{3.59}$$

Naturally, the consistency and boundedness results gained in the previous section transfer from the discrete bilinear forms to the discrete operators.

Proposition 3.28 *(Properties of discrete operators). The discrete operators A_h^{cf}, A_h^{upw} and S_h satisfy the following properties:*

i) Consistency*, i. e. for the exact solution $u \in V_\star$ of (1.35) it holds*

$$A_h^{\mathrm{cf}} u = \pi_h A u, \qquad A_h^{\mathrm{upw}} u = \pi_h A u. \tag{3.60}$$

Indeed, equation (3.60) holds true for all functions $v \in V_\star$. In addition, the stabilization operator S_h satisfies $S_h v = 0$ for all $v \in V_\star$.

ii) Boundedness*, i. e. for all $v \in V_{\star h}$ it holds*

$$\|A_h^{\mathrm{cf}} v\|_V \le c_\infty \|\nabla_h \times v\|_V + C_{\mathrm{S}} h^{-1/2} |v|_S, \tag{3.61}$$

$$\|A_h^{\mathrm{upw}} v\|_V \le c_\infty \|\nabla_h \times v\|_V + C'_{\mathrm{S}} h^{-1/2} |v|_S, \tag{3.62}$$

$$\|S_h v\|_V \le C_{\mathrm{S}} h^{-1/2} |v|_S, \tag{3.63}$$

with $C_{\mathrm{S}} = C_{\mathrm{bnd}} c_\infty^{1/2}$ and $C'_{\mathrm{S}} = 2 C_{\mathrm{bnd}} c_\infty^{1/2}$.

iii) Skew-adjointness *of the operator A_h^{cf} on V_h, i. e. for all $v_h, \widehat{v}_h \in V_h$ it holds*

$$(A_h^{\mathrm{cf}} v_h, \widehat{v}_h)_V = -(A_h^{\mathrm{cf}} \widehat{v}_h, v_h)_V. \tag{3.64}$$

iv) Dissipativity *of the operator $-A_h^{\mathrm{upw}}$ on V_h, i. e. for all $v_h \in V_h$ it holds*

$$(-A_h^{\mathrm{upw}} v_h, v_h)_V = -|v_h|_S^2 \le 0. \tag{3.65}$$

Proof: i) We have proven in Lemmata 3.14 and 3.20 that there holds

$$(A_h^{\mathrm{cf}} u, \varphi_h)_V = a_h^{\mathrm{cf}}(u, \varphi_h) = a(u, \varphi_h) = (Au, \varphi_h)_V \qquad \forall \varphi_h \in V_h,$$

and

$$(A_h^{\mathrm{upw}} u, \varphi_h)_V = a_h^{\mathrm{upw}}(u, \varphi_h) = a(u, \varphi_h) = (Au, \varphi_h)_V \qquad \forall \varphi_h \in V_h.$$

Hence, (3.60) follows with (3.58). The same arguments apply for all $v \in V_\star$. The assertion $S_h v = 0$ for all $v \in V_\star$ is seen with (3.6) and (3.7).
ii) For $v \in V_{\star h}$ it holds

$$\begin{aligned}\|A_h^{\mathrm{cf}} v\|_V &= \sup_{\substack{\varphi_h \in V_h \\ \|\varphi_h\|_V = 1}} |(A_h^{\mathrm{cf}} v, \varphi_h)_V| = \sup_{\substack{\varphi_h \in V_h \\ \|\varphi_h\|_V = 1}} |a_h^{\mathrm{cf}}(v, \varphi_h)| \\ &\le \sup_{\substack{\varphi_h \in V_h \\ \|\varphi_h\|_V = 1}} \left[(c_\infty \|\nabla_h \times v\|_V + C_{\mathrm{bnd}} c_\infty^{-1/2} h^{-1/2} |v|_S) \|\varphi_h\|_V \right] \\ &= c_\infty \|\nabla_h \times v\|_V + C_{\mathrm{S}} h^{-1/2} |v|_S,\end{aligned}$$

where we used the boundedness of a_h^{cf}, see (3.48). The boundedness for A_h^{upw} and S_h are proven analogously.
iii) The skew-adjointness of A_h^{cf} follows directly by Lemma 3.14.
iv) The dissipative property of $-A_h^{\mathrm{upw}}$ is seen by Lemma 3.20. □

The discretizations in operator form read: We search for $u_h^{\mathrm{cf}}, u_h^{\mathrm{upw}} \in C^1(0, T; V_h)$ such that there holds

$$\partial_t u_h^{\mathrm{cf}} + A_h^{\mathrm{cf}} u_h^{\mathrm{cf}} = \pi_h g, \tag{3.66}$$

and

$$\partial_t u_h^{\mathrm{upw}} + A_h^{\mathrm{upw}} u_h^{\mathrm{upw}} = \pi_h g. \tag{3.67}$$

We use the projection of u_0 as initial value, i. e. we require $u_h^{\mathrm{cf}}(0) = u_h^{\mathrm{upw}}(0) = \pi_h u_0$.

3.5 Stability

The following theorem reveals that the discrete schemes (3.66) and (3.67) are stable in the same sense as the continuous problem, see Theorem 1.19.

Theorem 3.29 *(Stability of discrete schemes). Let $u_h^{\mathrm{cf}} \in V_h$ be the solution of (3.66) and $u_h^{\mathrm{upw}} \in V_h$ be the solution of (3.67). Then, for all $t \in [0, T]$ the following stability results hold:*

i) In the homogeneous case*, i. e. for $g \equiv 0$, we have*

$$\|u_h^{\mathrm{cf}}(t)\|_V = \|\pi_h u_0\|_V, \tag{3.68}$$

and

$$\|u_h^{\mathrm{upw}}(t)\|_V^2 + 2 \int_0^t |u_h^{\mathrm{upw}}(s)|_S^2 \, ds = \|\pi_h u_0\|_V^2. \tag{3.69}$$

ii) In the inhomogeneous case*, we have*

$$\|u_h^{\mathrm{cf}}(t)\|_V^2 \le C_0 \left(\|u_0\|_V^2 + T \int_0^t \|g(s)\|_V^2 \, ds \right), \tag{3.70}$$

and

$$\|u_h^{\mathrm{upw}}(t)\|_V^2 + 2 \int_0^t |u_h^{\mathrm{upw}}(s)|_S \, ds \le C_0 \left(\|u_0\|_V^2 + T \int_0^t \|g(s)\|_V^2 \, ds \right), \tag{3.71}$$

with $C_0 := e$.

Let us state an important remark before proving this theorem.

Remark 3.30 (Energy conservation versus dissipation). Recall that the V-norm can be associated with the energy of the Maxwell system and that the continuous solution is conservative, i. e. $\|u(t)\|_V = \|u_0\|_V$ for all $t \geq 0$. We conclude from (3.68) that the centered fluxes discretization preserves this property whereas the upwind fluxes discretization decreases the energy, see (3.69), and thus is *dissipative*. Indeed, revisiting the definition of the S-seminorm, we see that the amount of dissipation is related to the norm of the tangential jumps of the solution. We have already commented that this quantity is related to the quality of the discrete solution. Thus, we expect that a high-order solutions, i. e. solutions obtained with high polynomial degree k and small meshsize h, admit less dissipation than low-order solutions. Furthermore, we will see that the appearance of the S-seminorm on the LHS of (3.71) allows us to prove a better convergence result rather than in the centered fluxes case. ◇

Proof: We multiply (3.66) by u_h^{cf} and (3.67) by u_h^{upw} which yields

$$(\partial_t u_h^{\mathrm{cf}}, u_h^{\mathrm{cf}})_V + (A_h^{\mathrm{cf}} u_h^{\mathrm{cf}}, u_h^{\mathrm{cf}})_V = (\pi_h g, u_h^{\mathrm{cf}})_V,$$

and

$$(\partial_t u_h^{\mathrm{upw}}, u_h^{\mathrm{upw}})_V + (A_h^{\mathrm{upw}} u_h^{\mathrm{upw}}, u_h^{\mathrm{upw}})_V = (\pi_h g, u_h^{\mathrm{upw}})_V.$$

Using the identity $(\partial_t v, v)_V = \frac{1}{2}\frac{d}{dt}\|v\|_V^2$ together with the skew-adjointness of A_h^{cf} in the first equation and the dissipative property of $-A_h^{\mathrm{upw}}$ in the second equation yields

$$\frac{d}{dt}\|u_h^{\mathrm{cf}}\|_V^2 = 2(\pi_h g, u_h^{\mathrm{cf}})_V, \tag{3.72}$$

and

$$\frac{d}{dt}\|u_h^{\mathrm{upw}}\|_V^2 + 2|u_h^{\mathrm{upw}}|_S^2 = 2(\pi_h g, u_h^{\mathrm{upw}})_V. \tag{3.73}$$

i) Obviously, for $g \equiv 0$, assertions (3.68) and (3.69) follow by integrating (3.72) and (3.73) from 0 to t.
ii) In the inhomogeneous case we proceed by applying the weighted Young's inequality A.3 with $\gamma = T$ to (3.72) yielding

$$\frac{d}{dt}\|u_h^{\mathrm{cf}}(t)\|_V^2 \leq T\|g(t)\|_V^2 + \frac{1}{T}\|u_h^{\mathrm{cf}}(t)\|_V^2,$$

where we further used (3.59). Integrating from 0 to t gives the inequality

$$\|u_h^{\mathrm{cf}}(t)\|_V^2 \leq \|u_h^{\mathrm{cf}}(0)\|_V^2 + T\int_0^t \|g(s)\|_V^2\, ds + \frac{1}{T}\int_0^t \|u(s)\|_V^2\, ds,$$

which is of a form to which the continuous Gronwall lemma A.4 applies. Thus, we infer

$$\|u_h^{\mathrm{cf}}(t)\|_V^2 \leq e^{t/T}\left(\|u_h^{\mathrm{cf}}(0)\|_V^2 + T\int_0^t \|g(s)\|_V^2\, ds\right).$$

Clearly, there holds $e^{t/T} \leq e$ for $t \in [0,T]$ and by (3.59) we see $\|u_h^{\mathrm{cf}}(0)\|_V \leq \|u_0\|_V$. Hence, (3.70) is proven. It remains to prove (3.71). Therefore, we apply the weighted Young's inequality A.3 with $\gamma = T$ to (3.73) and integrate from 0 to T. This gives

$$\|u_h^{\mathrm{upw}}(t)\|_V^2 + 2\int_0^t |u_h^{\mathrm{upw}}(s)|_S^2\, ds \leq \|u_0\|_V^2 + T\int_0^t \|g(s)\|_V^2\, ds + \frac{1}{T}\int_0^t \|u_h^{\mathrm{upw}}(s)\|_V^2\, ds. \tag{3.74}$$

Furthermore, from the continuous Gronwall lemma A.4 we conclude

$$\|u_h^{\mathrm{upw}}(t)\|_V^2 \leq e^{t/T}\left(\|u_0\|_V^2 + T\int_0^t \|g(s)\|_V^2\,ds\right).$$

Plugging this into (3.74) yields

$$\begin{aligned}\|u_h^{\mathrm{upw}}(t)\|_V^2 + 2\int_0^t |u_h^{\mathrm{upw}}(s)|_S^2\,ds \leq& \|u_0\|_V^2 + T\int_0^t \|g(s)\|_V^2\,ds \\ &+ \frac{1}{T}\int_0^t e^{s/T}\left(\|u_0\|_V^2 + T\int_0^s \|g(r)\|_V^2\,dr\right)ds. \end{aligned} \tag{3.75}$$

Since $\|g(r)\|_V^2$ is non-negative and $s \in (0,t)$ we can estimate the last integral by

$$\begin{aligned}\frac{1}{T}\int_0^t e^{s/T}\left(\|u_0\|_V^2 + T\int_0^s \|g(r)\|_V^2\,dr\right)ds \leq& \frac{1}{T}\int_0^t e^{s/T}\left(\|u_0\|_V^2 + T\int_0^t \|g(r)\|_V^2\,ds\right)ds \\ =& \frac{1}{T}\left(\|u_0\|_V^2 + T\int_0^t \|g(r)\|_V^2\,dr\right)Te^{s/T}\Big|_{s=0}^t \\ =& \left(e^{t/T}-1\right)\left(\|u_0\|_V^2 + T\int_0^t \|g(r)\|_V^2\,dr\right).\end{aligned}$$

Inserting this into (3.75) proves the assertion. □

3.6 Convergence

In the following we use the notation

$$V_{\star,k+1} := \mathcal{D}(A) \cap H^{k+1}(\mathcal{T}_h)^6.$$

We begin the convergence analysis by investigating the types of errors appearing in the discretizations (3.66) and (3.67).

3.6.1 Error Analysis

Definition 3.31 (Error types). Let $u \in V_\star$ denote the exact solution of (1.35) and $u_h^{\mathrm{cf}}, u_h^{\mathrm{upw}} \in V_h$ denote the discrete solutions of (3.66) and (3.67), respectively. We define the *spatial discretization errors*

$$e^{\mathrm{cf}}(t) := u(t) - u_h^{\mathrm{cf}}(t), \qquad e^{\mathrm{upw}}(t) := u(t) - u_h^{\mathrm{upw}}(t).$$

Furthermore, we split the errors into two parts

$$e^{\mathrm{cf}}(t) = e_\pi(t) - e_h^{\mathrm{cf}}(t), \qquad e^{\mathrm{upw}}(t) = e_\pi(t) - e_h^{\mathrm{upw}}(t),$$

where $e_\pi(t)$ is the *projection error*

$$e_\pi(t) := u(t) - \pi_h u(t),$$

and $e_h^{\mathrm{cf}}(t)$, $e_h^{\mathrm{upw}}(t)$ are given as

$$e_h^{\mathrm{cf}}(t) := u_h^{\mathrm{cf}}(t) - \pi_h u(t), \qquad e_h^{\mathrm{upw}}(t) := u_h^{\mathrm{upw}}(t) - \pi_h u(t).$$

We recall that by Definition 3.27 of the V-projection there holds

$$(e_\pi(t), \varphi_h)_V = 0 \qquad \forall \varphi_h \in V_h. \tag{3.76}$$

The projection error e_π arises from replacing the continuous space $V_\star$ by the finite space V_h. Indeed, $\pi_h u$ is the best-approximation to u in V_h, and thus the projection error e_π is the minimal error we can obtain. The splitting of the error provides two advantages. First, from Section 2.2.3 we already have bounds for the projection error. They are stated in the following Lemma 3.32. Secondly, the errors e_h^{cf} and e_h^{upw} measure the error between the discrete solutions u_h^{cf}, u_h^{upw} and the best-approximation $\pi_h u$. All three terms are elements of V_h and consequently the errors are in V_h, too. This allows us to state discrete error equations in V_h which are given in Lemma 3.33.

Lemma 3.32 *(Bounds for the projection error). Let $v \in H^{k+1}(\mathcal{T}_h)^6$. Then, the projection error is bounded by*

$$\|v - \pi_h v\|_V \le C_\pi h^{k+1} |v|_{H^{k+1}(\mathcal{T}_h)^6},$$

and its broken curl by

$$\|\nabla_h \times (v - \pi_h v)\|_V \le C_\pi h^k |v|_{H^{k+1}(\mathcal{T}_h)^6}.$$

The constant is given by $C_\pi := C'_{\mathrm{app}} \max\{\mu_\infty^{1/2}, \varepsilon_\infty^{1/2}\}$ and is independent of the meshsize h.

Proof: Let $v = [H, E]^T \in H^{k+1}(\mathcal{T}_h)^6$ and set $\xi_\pi = v - \pi_h v$, i. e.

$$\xi_\pi = \begin{bmatrix} \xi_\pi^H \\ \xi_\pi^E \end{bmatrix} = \begin{bmatrix} H - \pi_h H \\ E - \pi_h E \end{bmatrix}.$$

Then, there holds

$$\|\xi_\pi\|_V = \left\| \begin{bmatrix} \mu^{1/2} \xi_\pi^H \\ \varepsilon^{1/2} \xi_\pi^E \end{bmatrix} \right\|_{L^2(\Omega)^6} \le \max\{\mu_\infty^{1/2}, \varepsilon_\infty^{1/2}\} \|\xi_\pi\|_{L^2(\Omega)^6}.$$

Applying Lemma 2.23 on each mesh element K yields

$$\|\xi_\pi\|_V \le C'_{\mathrm{app}} \max\{\mu_\infty^{1/2}, \varepsilon_\infty^{1/2}\} h^{k+1} |v|_{H^{k+1}(\mathcal{T}_h)^6}.$$

Hence, the first assertion follows. For the second assertion note that $\|\nabla_h \times \xi_\pi\|_{L^2(\Omega)^6} \le |\xi_\pi|_{H^1(\mathcal{T}_h)^6}$ and thus

$$\|\nabla_h \times \xi_\pi\|_V \le \max\{\mu_\infty^{1/2}, \varepsilon_\infty^{1/2}\} |\xi_\pi|_{H^1(\mathcal{T}_h)^6}.$$

Lemma 2.23 then yields the result. □

Clearly, if the exact solution satisfies $u \in V_{\star,k+1}$, Lemma 3.32 provides the bounds

$$\|e_\pi\|_V \le C_\pi h^{k+1} |u|_{H^{k+1}(\mathcal{T}_h)^6}, \tag{3.77}$$

and

$$\|\nabla_h \times e_\pi\|_V \le C_\pi h^k |u|_{H^{k+1}(\mathcal{T}_h)^6}. \tag{3.78}$$

Lemma 3.33 *(Error equations). For the errors e_h^{cf} and e_h^{upw} there hold following discrete evolution equations,*

$$\partial_t e_h^{\mathrm{cf}} + A_h^{\mathrm{cf}} e_h^{\mathrm{cf}} = A_h^{\mathrm{cf}} e_\pi, \qquad e_h^{\mathrm{cf}}(0) = 0, \tag{3.79}$$

and

$$\partial_t e_h^{\mathrm{upw}} + A_h^{\mathrm{upw}} e_h^{\mathrm{upw}} = A_h^{\mathrm{upw}} e_\pi, \qquad e_h^{\mathrm{upw}}(0) = 0. \tag{3.80}$$

Proof: Projecting the continuous problem (1.35) onto V_h gives

$$\partial_t \pi_h u + \pi_h A u = \pi_h g, \qquad \pi_h u(0) = \pi_h u_0,$$

where we used that the projection operator is independet of the time t and therefore it holds $\partial_t \pi_h = \pi_h \partial_t$. Owing to the consistency of the operator A_h^{cf}, see Proposition 3.28, this is equivalent to

$$\partial_t \pi_h u + A_h^{\mathrm{cf}} u = \pi_h g. \tag{3.81}$$

For the discrete solution u_h^{cf} there holds $u_h^{\mathrm{cf}}(0) = \pi_h u_0$ as well as

$$\partial_t u_h^{\mathrm{cf}} + A_h^{\mathrm{cf}} u_h^{\mathrm{cf}} = \pi_h g. \tag{3.82}$$

Clearly, there holds $e_h^{\mathrm{cf}}(0) = 0$ and by substracting (3.81) from (3.82) we see

$$\partial_t e_h^{\mathrm{cf}} - A_h^{\mathrm{cf}} e = 0.$$

Hence, (3.79) follows by the splitting of the error, i. e. by $e = e_\pi - e_h^{\mathrm{cf}}$. Assertion (3.80) is proven analogously. □

Combining the error equation (3.79) with the stability result we can prove the convergence of the centered fluxes discretization.

Theorem 3.34 *(Convergence for centered fluxes). Let $u \in C^1(0,T;V) \cap C(0,T;V_{\star,k+1})$ be the exact solution of (1.35) and $u_h^{\mathrm{cf}} \in C^1(0,T;V_h)$ be the discrete solution of (3.66). Then, for the error there holds*

$$\|e^{\mathrm{cf}}(t)\|_V^2 \le C_{\mathrm{cf}} T h^{2k} \int_0^t |u(s)|^2_{H^{k+1}(\mathcal{T}_h)^6}\, ds + C'_{\mathrm{cf}} h^{2k+2} |u(t)|^2_{H^{k+1}(\mathcal{T}_h)^6}, \tag{3.83}$$

with $C_{\mathrm{cf}} = 2C_0 C_\pi^2 (c_\infty + C_{\mathrm{S}}^2)^2$ and $C'_{\mathrm{cf}} = 2C_\pi^2$ both independent of h.

Proof: We apply the stability result for the centered flux scheme (3.70) to the error equation (3.79) and obtain

$$\|e_h^{\mathrm{cf}}(t)\|_V^2 \le C_0 T \int_0^t \|A_h^{\mathrm{cf}} e_\pi(s)\|_V^2\, ds.$$

The boundedness of the operator A_h^{cf} (3.61) and the bound of the S-seminorm (3.56) yield

$$\begin{aligned}\|e_h^{\mathrm{cf}}(t)\|_V^2 \le& C_0 T \int_0^t (c_\infty \|\nabla_h \times e_\pi(s)\|_V + C_{\mathrm{S}} h^{-1/2} |e_\pi(s)|_S)^2\, ds \\ \le& C_0 T \int_0^t (c_\infty \|\nabla_h \times e_\pi(s)\|_V + C_{\mathrm{S}}^2 h^{-1} \|e_\pi(s)\|_V)^2\, ds.\end{aligned}$$

Next, we use the bounds on the projection errors (3.77) and (3.78) to infer

$$\|e_h^{\mathrm{cf}}(t)\|_V^2 \le C_0 C_\pi^2 (c_\infty + C_{\mathrm{S}}^2)^2 h^{2k} T \int_0^t |u(s)|^2_{H^{k+1}(\mathcal{T}_h)^6}\, ds. \tag{3.84}$$

Young's inequality yields for the full error

$$\|e^{\mathrm{cf}}(t)\|_V^2 \le 2\|e_h^{\mathrm{cf}}(t)\|_V^2 + 2\|e_\pi(t)\|_V^2 \le 2\|e_h^{\mathrm{cf}}(t)\|_V^2 + 2C_\pi^2 h^{2k+2} |u(t)|^2_{H^{k+1}(\mathcal{T}_h)^6}, \tag{3.85}$$

where we used (3.77) in the second inequality. Combining (3.84) and (3.85) yields the assertion. □

This theorem establishes convergence of the suboptimal order h^k for the centered fluxes scheme. Clearly, this result also holds true for the upwind case since the three ingredients, i. e. stability of the scheme, an error equation in form of the evolution problem and the bounds for the projection errors, apply analogously in this case. The crucial difference in the upwind case is that we can do better. Therefore, we essentially need the following property. For completeness we state it also for the centered fluxes operator but we will see that this does not improve the convergence result in this case.

Lemma 3.35 *For the projection error there holds*

$$|(A_h^{\mathrm{cf}} e_\pi, \varphi_h)_V| \leq C'_\pi h^{-1/2} |\varphi_h|_S \|e_\pi\|_V \qquad \forall \varphi_h \in V_h, \tag{3.86}$$

in the centered fluxes case and

$$|(A_h^{\mathrm{upw}} e_\pi, \varphi_h)_V| \leq C''_\pi h^{-1/2} |\varphi_h|_S \|e_\pi\|_V \qquad \forall \varphi_h \in V_h \tag{3.87}$$

in the upwind fluxes case. The constants are given as $C'_\pi = (\sqrt{2} N_\partial^{1/2} + 1) C_{\mathrm{tr}} C_{\mathrm{qu}}^{-1/2} c_\infty^{1/2}$ *and* $C''_\pi = C'_\pi + C_{\mathrm{S}}$.

Proof: Let $e_\pi = [e_\pi^{\mathcal{H}}, e_\pi^{\mathcal{E}}]^T$ denote the projection error. We begin with the centered fluxes result. Using the partial integration form (3.34) we have for all $\varphi_h = [\phi_h, \psi_h]^T \in V_h$,

$$\begin{aligned}
(A_h^{\mathrm{cf}} e_\pi, \varphi_h)_V =& a_h^{\mathrm{cf}}(e_\pi, \varphi_h) \\
=& (e_\pi^{\mathcal{E}}, \nabla_h \times \phi_h)_{L^2(\Omega)^3} - (e_\pi^{\mathcal{H}}, \nabla_h \times \psi_h)_{L^2(\Omega)^3} \\
&+ \sum_{F \in \mathcal{F}_h^i} \left[(\{\!\{ e_\pi^{\mathcal{E}} \}\!\}_F^{c\varepsilon}, n_F \times [\![\phi_h]\!]_F)_{L^2(F)^3} - (\{\!\{ e_\pi^{\mathcal{H}} \}\!\}_F^{c\mu}, n_F \times [\![\psi_h]\!]_F)_{L^2(F)^3} \right] \\
&+ \sum_{F \in \mathcal{F}_h^b} (e_\pi^{\mathcal{H}}, n \times \psi_h)_{L^2(\Omega)^3}.
\end{aligned}$$

Since $\nabla_h \times \phi_h$, $\nabla_h \times \psi_h$ are elements of V_h the first to terms vanish due to (3.76). In order to bound the remaining terms we can use the same computations as in the proof of Theorem 3.22. Indeed we get by (3.31) and (3.52)

$$\begin{aligned}
&\sum_{F \in \mathcal{F}_h^i} \left[(\{\!\{ e_\pi^{\mathcal{E}} \}\!\}_F^{c\varepsilon}, n_F \times [\![\phi_h]\!]_F)_{L^2(F)^3} - (\{\!\{ e_\pi^{\mathcal{H}} \}\!\}_F^{c\mu}, n_F \times [\![\psi_h]\!]_F)_{L^2(F)^3} \right] \\
&= \sum_{F \in \mathcal{F}_h^i} \left[(n_F \times [\![\phi_h]\!]_F, \{\!\{ e_\pi^{\mathcal{E}} \}\!\}_F^{\overline{c\varepsilon}})_{L^2(F)^3} - (n_F \times [\![\psi_h]\!]_F, \{\!\{ e_\pi^{\mathcal{H}} \}\!\}_F^{\overline{c\mu}})_{L^2(F)^3} \right] \\
&\leq \widetilde{C} h^{-1/2} |\varphi_h|_S \|e_\pi\|_V,
\end{aligned}$$

with $\widetilde{C} = \sqrt{2} C_{\mathrm{tr}} C_{\mathrm{qu}}^{-1/2} c_\infty^{1/2} N_\partial^{1/2}$. Furthermore, by (3.53) we have

$$\sum_{F \in \mathcal{F}_h^b} (e_\pi^{\mathcal{H}}, n \times \psi_h)_{L^2(\Omega)^3} \leq \widetilde{C}' h^{-1/2} |\varphi_h|_S \|e_\pi\|_V,$$

where $\widetilde{C}' = C_{\mathrm{tr}} C_{\mathrm{qu}}^{-1/2} c_\infty^{1/2}$. This proves (3.86). For the upwind case we use the symmetry of the stabilization bilinear form on $V_{\star h}$ together with the stability result (3.54) to infer

$$|(S_h e_\pi, \varphi_h)_V| = |s_h(\varphi_h, e_\pi)| \leq C_{\mathrm{S}} h^{-1/2} |\varphi_h|_S \|e_\pi\|_V.$$

Assertion (3.87) then follows from (3.86) and the just shown bound. □

Theorem 3.36 *(Convergence for upwind fluxes). Let $u \in C^1(0,T;V) \cap C(0,T;V_{\star,k+1})$ be the exact solution of (1.35) and $u_h^{\mathrm{upw}} \in C^1(0,T;V_h)$ be the discrete solution of the upwind discretization (3.67). Then, for the error there holds*

$$\|e^{\mathrm{upw}}(t)\|_V^2 + \int_0^t |u_h(s)|_S^2 \, ds \\ \leq C_{\mathrm{upw}} h^{2k+1} \int_0^t |u(s)|^2_{H^{k+1}(\mathcal{T}_h)^6} \, ds + C'_{\mathrm{upw}} h^{2k+2} |u(t)|^2_{H^{k+1}(\mathcal{T}_h)^6},$$

with constants $C_{\mathrm{upw}} := 2((C''_\pi)^2 + C_{\mathrm{S}}^2)$ and $C'_{\mathrm{upw}} := 2C_\pi^2$.

Proof: We begin by multiplying the upwind error equation (3.80) by e_h^{upw},

$$\frac{d}{dt}\|e_h^{\mathrm{upw}}(t)\|_V^2 + 2(A_h^{\mathrm{upw}} e_h^{\mathrm{upw}}, e_h^{\mathrm{upw}})_V = 2(A_h^{\mathrm{upw}} e_\pi, e_h^{\mathrm{upw}})_V.$$

Employing the dissipative property (3.65) of the operator A_h^{upw} and integrating from 0 to t yields

$$\|e_h^{\mathrm{upw}}(t)\|_V^2 + 2\int_0^t |e_h^{\mathrm{upw}}(s)|_S^2 \, ds = 2\int_0^t (A_h^{\mathrm{upw}} e_\pi(s), e_h^{\mathrm{upw}}(s))_V \, ds, \tag{3.88}$$

since $e_h^{\mathrm{upw}}(0) = 0$. By applying (3.87) from the previous lemma we get

$$\|e_h^{\mathrm{upw}}(t)\|_V^2 + 2\int_0^t |e_h^{\mathrm{upw}}(s)|_S^2 \, ds \leq 2\int_0^t C''_\pi h^{-1/2} |e_h^{\mathrm{upw}}(s)|_S \|e_\pi(s)\|_V \, ds. \tag{3.89}$$

Using Young's inequality we can estimate the RHS by

$$2\int_0^t C''_\pi h^{-1/2} |e_h^{\mathrm{upw}}(s)|_S \|e_\pi(s)\|_V \, ds \leq \int_0^t \left[(C''_\pi h^{-1/2} \|e_\pi(s)\|_V)^2 + |e_h^{\mathrm{upw}}(s)|_S^2 \right] ds.$$

Inserting this into (3.89) and canceling the integral over $|e_h^{\mathrm{upw}}(s)|_S^2$ with its counterpart on the LHS we get

$$\|e_h^{\mathrm{upw}}(t)\|_V^2 + \int_0^t |e_h^{\mathrm{upw}}(s)|_S^2 \, ds \leq (C''_\pi)^2 h^{-1} \int_0^t \|e_\pi(s)\|_V^2 \, ds. \tag{3.90}$$

In order to bound the full error recall that by (3.45) it holds $|u|_S = 0$. Thus, we have

$$|u_h|_S \leq |u - u_h|_S + |u|_S = |e^{\mathrm{upw}}|_S \leq |e_\pi|_S + |e_h^{\mathrm{upw}}|_S.$$

Consequently, we see by Young's inequality and the splitting of the error that there holds

$$\|e^{\mathrm{upw}}(t)\|_V^2 + \int_0^t |u_h(s)|_S^2 \, ds \leq 2\|e_\pi(t)\|_V^2 + 2\|e_h^{\mathrm{upw}}(t)\|_V^2 + 2\int_0^t \left[|e_\pi(s)|_S^2 + |e_h^{\mathrm{upw}}(s)|_S^2 \right] ds.$$

Inserting (3.90) yields

$$\|e^{\mathrm{upw}}(t)\|_V^2 + \int_0^t |u_h(s)|_S^2 \, ds \leq 2\|e_\pi(t)\|_V^2 + 2\int_0^t \left[|e_\pi(s)|_S^2 + (C''_\pi)^2 h^{-1} \|e_\pi(s)\|_V^2 \right] ds.$$

Corollary 3.24 further provides $|e_\pi(s)|_S^2 \leq C_{\mathrm{S}}^2 h^{-1} \|e_\pi\|_V^2$ whence we infer with Lemma 3.32

$$\begin{aligned}\|e^{\mathrm{upw}}(t)\|_V^2 + \int_0^t |u_h(s)|_S^2 \, ds \leq & 2C_\pi^2 h^{2k+2} |u(t)|_{H^{k+1}(\mathcal{T}_h)^6}^2 \\ & + 2(C_{\mathrm{S}}^2 + (C_\pi'')^2) h^{2k+1} \int_0^t |u(s)|_{H^{k+1}(\mathcal{T}_h)^6}^2 \, ds,\end{aligned}$$

and the proof is finished. □

Remark 3.37 Revisiting the current proof we observe that we cannot transfer it to the centered fluxes case. In particular, notice that the crucial deduction is from (3.88) to (3.90). Thereby, we needed two results. The first is the estimate $(A_h^{\mathrm{upw}} e_\pi, e_h^{\mathrm{upw}})_V \leq C h^{-1/2} |e_h^{\mathrm{upw}}|_S \|e_\pi\|_V$ which is provided by Lemma 3.35 and which holds also true for the centered fluxes operator. Secondly, we canceled the error term $|e_h^{\mathrm{upw}}|_S$ on the RHS with its counterpart on the LHS which is provided by the dissipative property of the upwind operator, $(A_h^{\mathrm{upw}} e_h^{\mathrm{upw}}, e_h^{\mathrm{upw}})_V = |e_h^{\mathrm{upw}}|_S^2$. It is exactly the absence of this property which disables this proof for the centered fluxes case. We could continue by estimating the term $|e_h^{\mathrm{cf}}|_S$ instead of balancing it, but this would give the same (suboptimal) convergence rate as already provided by Theorem 3.34. In fact, our later numeircal results show that we cannot do better than this convergence rate. ◇

Chapter 4

Full Discretization

In Chapter 3 we have discretized Maxwell's equations (1.35) in space using dG methods. This led to a discrete evolution equation of the form

$$\partial_t u_h(t) = -A_h u_h(t) + g_h(t), \qquad t \in (0,T), \tag{4.1a}$$

$$u_h(0) = \pi_h u_0. \tag{4.1b}$$

posed in the finite dimensional space V_h. Here $A_h \in \{A_h^{\mathrm{cf}}, A_h^{\mathrm{upw}}\}$ is the discrete centered fluxes operator or the discrete upwind fluxes operator, see Definition 3.26, and we set $g_h = \pi_h g$. We have proven that the discrete scheme (4.1) is consistent and stable and furtheron that its solution $u_h(t)$ converges to the exact solution $u(t)$ as the meshsize tends to zero with convergence rate h^k in the centered fluxes case and with rate $h^{k+1/2}$ in the upwind fluxes case.

In this chapter we discretize the semi-discrete problem (4.1) in time with explicit RK methods. We point out that Burman, Ern and Fernández have proven in [2] the convergence for two- and three-stage RK methods for first-order differential equations of Friedrich's type and that this framework covers Maxwell's equations (1.35) and the associated dG semi-discretization (4.1). Our analysis is strongly motivated by this paper with the difference that we start with the forward Euler method rather than with two-stage RK methods. It is instructive to consider the forward Euler method, too, since the used techniques and the gained results resemble those for two-stage RK methods. Furthermore, we incorporate a stability analysis for the RK approximation of (4.1) for the case with no source term, i. e. $g_h \equiv 0$, and for the general case with source term.

This chapter is organised as follows: First we show that the discretized Maxwell operator A_h is be bounded on the discrete space V_h by $\mathcal{O}(h^{-1})$. Then, we introduce (explicit) RK methods and afterwards head towards proving the so called energy identities for the RK approximations. This identities enable us to prove stability. Next, we show that the error satisfies the same recursion as the RK approximation and deduce the convergence of order $h^k + \tau^s$ by the stability result, where s is the stage number of the RK method. Last, we improve this result to $h^{k+1/2} + \tau^s$ for the upwind fluxes case. This proof cannot be done relying on the stability results but we need to start from the energy identities.

4.1 Boundedness of A_h on V_h

The following result is crucial for the subsequent convergence analyis.

Theorem 4.1 *(Boundedness of A_h on V_h). Let $A_h \in \{A_h^{\mathrm{cf}}, A_h^{\mathrm{upw}}\}$. Then, for all $v_h \in V_h$ there holds,*

$$\|A_h v_h\|_V \le C_h c_\infty h^{-1} \|v_h\|_V, \tag{4.2}$$

with the associated constant $C_h \in \{C_h^{\mathrm{cf}}, C_h^{\mathrm{upw}}\}$ given as $C_h^{\mathrm{cf}} = C_{\mathrm{inv}} + C_{\mathrm{bnd}}^2$ and $C_h^{\mathrm{upw}} = C_{\mathrm{inv}} + 2C_{\mathrm{bnd}}^2$.

Proof: In Proposition 3.28 and Corollary 3.24 we have shown that for all $v_h \in V_h$ there holds

$$\|A_h^{\mathrm{cf}} v_h\|_V \leq c_\infty \|\nabla_h \times v_h\|_V + C_{\mathrm{S}}^2 h^{-1} \|v_h\|_V, \tag{4.3}$$

and

$$\|A_h^{\mathrm{upw}} v_h\|_V \leq c_\infty \|\nabla_h \times v_h\|_V + C_{\mathrm{S}}' C_{\mathrm{S}} h^{-1} \|v_h\|_V. \tag{4.4}$$

Recalling that we defined $C_{\mathrm{S}} = C_{\mathrm{bnd}} c_\infty^{1/2}$ and $C_{\mathrm{S}}' = 2 C_{\mathrm{bnd}} c_\infty^{1/2}$ we see that the second terms in (4.3) and (4.4) meet the bound (4.2). Consequently, it remains to estimate the curl term $\|\nabla_h \times v_h\|_V$. According to the definition of the broken curl and the V-norm we have for all $v_h = [H_h, E_h]^T \in V_h$,

$$\begin{aligned} \|\nabla_h \times v_h\|_V^2 &= \sum_{K \in \mathcal{T}_h} \|\nabla \times v_h\|_{V(K)}^2 = \sum_{K \in \mathcal{T}_h} \left\| \begin{bmatrix} \mu_K^{1/2} \nabla \times H_h \\ \varepsilon_K^{1/2} \nabla \times E_h \end{bmatrix} \right\|_{L^2(K)^6}^2 \\ &= \sum_{K \in \mathcal{T}_h} \left[\mu_K \|\nabla \times H_h\|_{L^2(K)^3}^2 + \varepsilon_K \|\nabla \times E_h\|_{L^2(K)^3}^2 \right]. \end{aligned} \tag{4.5}$$

Furthermore, there holds

$$\|\nabla \times H_h\|_{L^2(K)^3} \leq |H_h|_{H^1(K)^3} = \|\nabla H_h\|_{L^2(K)^{3\times 3}} \leq C_{\mathrm{inv}}^2 h_K^{-2} \|H_h\|_{L^2(K)^3},$$

where we applied the inverse inequality (2.13) componentwise in the last estimate. Analogously, we have

$$\|\nabla \times E_h\|_{L^2(K)^3}^2 \leq C_{\mathrm{inv}}^2 h_K^{-2} \|E_h\|_{L^2(K)^3}^2.$$

Inserting this estimates in (4.5) gives

$$\begin{aligned} \|\nabla_h \times v_h\|_V^2 \leq & C_{\mathrm{inv}}^2 \sum_{K \in \mathcal{T}_h} h_K^{-2} \left[\mu_K \|H_h\|_{L^2(K)^3}^2 + \varepsilon_K^2 \|E_h\|_{L^2(K)^3}^2 \right] \\ = & C_{\mathrm{inv}}^2 \sum_{K \in \mathcal{T}_h} h_K^{-2} \|v_h\|_{V(K)}^2 \leq C_{\mathrm{inv}}^2 h^{-2} \sum_{K \in \mathcal{T}_h} \|v_h\|_{V(K)}^2 = C_{\mathrm{inv}}^2 h^{-2} \|v_h\|_V^2, \end{aligned}$$

where we used Assumption 3.21 in the last inequality. □

Now, we construct the time discretization of (4.1) with RK methods.

4.2 Runge-Kutta Methods

We begin by discretizing the time interval $[0, T]$ by a discrete set $\{t_n\}_{n=1}^N$ on which we want to approximate the semi-discrete solution $u_h^n \approx u_h(t_n)$. For simplicity we assume that the points $\{t_n\}$ are equidistant, i. e. $t_{n+1} - t_n = \tau$ for all $n = 1, \ldots, N$. We call τ the *step size* and we assume that the finaltime T is a multiple of the step size, i. e. there is an integer $N \in \mathbb{N}$ such that $T = N\tau$.

RK methods start with a given intial value $u_h^0 = u_h(0)$ and compute consecutively a sequence of approximations $\left\{u_h^n\right\}_n$ using the approximation from the previous step only to compute the current approximation: $u_h^n \rightsquigarrow u_h^{n+1}$. This characterizes RK methods as *single step methods*.

4.2.1 Construction of Runge-Kutta Methods

Let us at first notice that we construct RK methods in this thesis for linear equations (4.1) only. However, the construction stays exactly the same for a general equation $y'(t) = g(t, y(t))$, but we skip it since we do not need this case.

The construction of RK methods relies on the following integral representation of the solution of the evolution equation (4.1)

$$u_h(t_{n+1}) = u_h(t_n) + \int_{t_n}^{t_{n+1}} \left(-A_h u_h(s) + g_h(s)\right) ds.$$

The main idea is to approximate the integral by a quadrature rule. Usually, quadrature is performed in the refernce interval $[0,1]$ and thus we transform the integral from above to this interval,

$$u_h(t_{n+1}) = u_h(t_n) + \tau \int_0^1 \left(-A_h u_h(t_n + \tau s) + g_h(t_n + \tau s)\right) ds.$$

Now let $c_1, \dots, c_s \in [0,1]$ be the quadrature nodes and $b_1, \dots, b_s$ the associated weights. Then, the quadrature rule reads

$$\tau \int_0^1 \left(-A_h u_h(t_n + \tau s) + g_h(t_n + \tau s)\right) ds \approx \tau \sum_{i=1}^{s} b_i \left[-A_h u_h(t_n + c_i\tau) + g_h(t_n + c_i\tau)\right].$$

In order to construct a numerical method we have to approximate $u_h(t_n + c_i\tau)$, too. Clearly, it holds

$$u_h(t_n + c_i\tau) = u_h(t_n) + \tau \int_0^{c_i} \left(-A_h u_h(t_n + \tau s) + g_h(t_n + \tau s)\right) ds,$$

which leads to the idea of approximating $u_h(t_n + c_i\tau)$ by a quadrature rule again. It is meaningful to use the same quadrature nodes c_i as above. Furthermore, we choose weights a_{ij}, $i, j = 1, \dots, s$, such that the quadrature reads:

$$\tau \int_0^{c_i} \left(-A_h u_h(t_n + \tau s) + g_h(t_n + \tau s)\right) ds \approx \tau \sum_{j=1}^{s} a_{ij} \left[-A_h u_h(t_n + c_j\tau) + g_h(t_n + c_j\tau)\right].$$

Denoting with u_h^{ni} the approximations to the inner stages $u_h(t_n + c_i\tau)$ we can state a *general RK* method compactly as,

$$u_h^{ni} = u_h^n + \tau \sum_{j=1}^{s} a_{ij} \left[-A_h u_h^{nj} + g_h^{nj}\right], \quad i = 1, \dots, s, \tag{4.6a}$$

$$u_h^{n+1} = u_h^n + \tau \sum_{i=1}^{s} b_i \left[-A_h u_h^{ni} + g_h^{ni}\right], \tag{4.6b}$$

where we used the notation $g_h^{ni} := g_h(t_n + c_i\tau)$. The number s is called the *stage number* of the method. Usually, the coefficients of a RK scheme are collected in a so-called *Butcher tableau*

$$\left[\begin{array}{c|ccc} c_1 & a_{11} & \dots & a_{1s} \\ \vdots & \vdots & & \vdots \\ c_s & a_{s1} & \dots & a_{ss} \\ \hline & b_1 & \dots & b_s \end{array}\right].$$

We make the following assumption on the coefficients a_{ij}, b_i and c_i.

Assumption 4.2 *(Simplifying assumptions). We assume that the coefficients of the RK method satisfy the following conditions:*

$$\sum_{i=1}^{s} b_i = 1, \qquad \sum_{j=1}^{i-1} a_{ij} = c_i. \tag{4.7}$$

This assumption ensures that in the simple case of an ODE with constant right-hand side a RK method provides the exact solution at every time point t_n and at every inner stage $t_n + c_i\tau$, i. e. $u_h^n = u_h(t_n)$ and $u_h^{ni} = u_h(t_n + c_i\tau)$.

4.2.2 Explicit Runge-Kutta Methods

We see from (4.6a) that in general the inner stage u_h^{ni} can depend on all inner stages $u_h^{n1}, \dots, u_h^{ns}$. The constituting property of an *explicit Runge-Kutta method* is that we choose $a_{ij} = 0$ for $i \geq j$. Thus, every inner stage u_h^{ni} only depends on the stages $u_h^{n1}, \dots, u_h^{n(i-1)}$. which allows to compute the inner stages recursively.

Definition 4.3 (Explicit Runge-Kutta method). A general explicit s-stage Runge-Kutta approximation of the evolution equation (4.1) is given by

$$u_h^{ni} = u_h^n + \tau \sum_{j=1}^{i-1} a_{ij} \left[-A_h u_h^{nj} + g_h^{nj} \right], \quad i = 1, \dots, s, \tag{4.8a}$$

$$u_h^{n+1} = u_h^n + \tau \sum_{i=1}^{s} b_i \left[-A_h u_h^{ni} + g_h^{ni} \right]. \tag{4.8b}$$

Henceforth, we consider only explicit RK methods with one, two or three stages and often refer to them as RK1, RK2 and RK3 methods. We begin by giving some examples for these methods.

4.2.3 Examples

Let us at first state the well-known *forward* or *explicit Euler method* which is a one-stage RK method given by

$$\begin{array}{c|c} 0 & 0 \\ \hline & 1 \end{array} \qquad \begin{cases} u_h^{n1} = u_h^n, \\ u_h^{n+1} = u_h^n - \tau A_h u_h^n + \tau g_h^n. \end{cases}$$

Two examples of RK2 methods are the *Runge method* which is defined by

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1/2 & 1/2 & 0 \\ \hline & 0 & 1 \end{array} \qquad \begin{cases} u_h^{n1} = u_h^n, \\ u_h^{n2} = u_h^n - \frac{1}{2}\tau A_h u_h^{n1} + \frac{1}{2}\tau g_h^n, \\ u_h^{n+1} = u_h^n - \tau A_h u_h^{n2} + \tau g_h^{n+1/2}, \end{cases}$$

where we used the notation $g_h^{n+1/2} := g_h(t_n + \frac{1}{2}\tau)$, and the *two-stage Heun method* defined by

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ \hline & 1/2 & 1/2 \end{array} \qquad \begin{cases} u_h^{n1} = u_h^n, \\ u_h^{n2} = u_h^n - \tau A_h u_h^{n1} + \tau g_h^n, \\ u_h^{n+1} = u_h^n - \frac{1}{2}\tau A_h u_h^{n1} + \frac{1}{2}\tau g_h^n - \frac{1}{2}\tau A_h u_h^{n2} + \frac{1}{2}\tau g_h^{n+1}. \end{cases}$$

Finally, an example for a RK3 scheme is the *three-stage Heun method* given by

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ 1/3 & 1/3 & 0 & 0 \\ 2/3 & 0 & 2/3 & 0 \\ \hline & 1/4 & 0 & 3/4 \end{array} \qquad \begin{cases} u_h^{n1} = u_h^n, \\ u_h^{n2} = u_h^n - \frac{1}{3}\tau A_h u_h^{n1} + \frac{1}{3}\tau g_h^n, \\ u_h^{n3} = u_h^n - \frac{2}{3}\tau A_h u_h^{n2} + \frac{2}{3}\tau g_h^{n+1/3}, \\ u_h^{n+1} = u_h^n - \frac{1}{4}\tau A_h u_h^{n1} + \frac{1}{4}\tau g_h^n - \frac{3}{4}\tau A_h u_h^{n3} + \frac{3}{4}\tau g_h^{n+2/3}, \end{cases}$$

with the same notation for the source-term as above.

4.2.4 Order Conditions

An important quantity in the analysis of RK methods is the *order* of the method (see e.g. the textbook [7]). In order to define it we consider the case of a general ODE $y'(t) = \widehat{g}(t, y(t))$. We say that a RK method is of order p if the approximation after one step, y^1, satisfies

$$\|y(t_1) - y^1\| \leq C\tau^{p+1},$$

given the exact initial value as a starting point, $y^0 = y(t_0)$, and that the RHS is smooth enough, i. e. $\widehat{g} \in C^{p+1}$. The order of a RK method can be examined systematically using rooted trees which yield certain *order conditions* on the coefficients (cf. [9, Section 8.6-8.7], [7, Section II.2]). Furthermore, it is shown in [9, Theorem 8.13] that an explicit s-stage RK method has at most order s. From now on, we restrict our consideration to explicit RK methods having this maximal order. This is guaranteed by the simplifying assumption (4.7) together with the following order conditions ([7, Theorem 2.13]).

Assumption 4.4 *(Order conditions). We assume that the coefficients of the RK2 methods satisfy*

$$b_2 c_2 = \frac{1}{2}, \tag{4.9}$$

and the coefficients of the RK3 methods satisfy

$$b_2 c_2 + b_3 c_3 = \frac{1}{2}, \qquad b_2 c_2^2 + b_3 c_3^2 = \frac{1}{3}, \qquad b_3 a_{32} c_2 = \frac{1}{6}. \tag{4.10}$$

Remark 4.5 All examples presented in Section 4.2.3 satisfy the Assumptions 4.2 and 4.4. Furthermore, the forward Euler method is the only explicit one-stage RK method satisfying the simplifying assumption (4.7). ◇

Due to Assumptions 4.2 and 4.4 we know that for an s-stage RK approximation (with $s \in \{1, 2, 3\}$) of the evolution equation (4.1) we have

$$\|u_h(t_1) - u_h^1\|_V \leq C(h)\tau^{s+1}, \tag{4.11}$$

if the source term satisfies $g_h \in C^{s+1}(0, T; V_h)$. From the local bound (4.11) we get a global error estimate of the form

$$\|u_h(t_{n+1}) - u_h^{n+1}\|_V \leq C(h)\tau^s, \tag{4.12}$$

see therefore [9, Theorem 8.5], [7, Theorem 3.6]. But, we point out that constants in this estimates depend on the operator norm $\|A_h\|_{V_h \leftarrow V_h}$ which is proportional to h^{-1}. Therefore, the bound (4.12) is only applicable if we consider the evolution equation for a fixed meshsize. Since we are interested in the convergence of the fully discrete scheme, i. e. convergence when τ and h both tend to zero, we have to use a different approach. Furthermore, this means that we deal with a problem which gets more and more ill-posed when the meshsize tends to zero making the analysis more complicated. We will therefore introduce an energy technique which is suitable for such problems.

We begin by stating the considered RK methods with eliminated stages.

Lemma 4.6 *(Explicit RK form without inner stages). Let $s \in \{1, 2, 3\}$ and let $\{u_h^n\}_n$ be an s-stage RK approximation of the evolution equation (4.1). Then, we have the following recursions:*

i) For RK1 methods:

$$u_h^{n+1} = u_h^n - \tau A_h u_h^n + \tau g_h^n. \tag{4.13}$$

ii) For RK2 methods:

$$u_h^{n+1} = u_h^n - \tau A_h u_h^n + \frac{1}{2}\tau^2 A_h^2 u_h^n + \tau b_1 g_h^n + \tau b_2 g_h^{n2} - \frac{1}{2}\tau^2 A_h g_h^n, \tag{4.14}$$

iii) For RK3 methods:

$$\begin{aligned} u_h^{n+1} =& u_h^n - \tau A_h u_h^n + \frac{1}{2}\tau^2 A_h^2 u_h^n - \frac{1}{6}\tau^3 A_h^3 u_h^n + \tau b_1 g_h^n + \tau b_2 g_h^{n2} + \tau b_3 g_h^{n3} \\ & - \tau^2 b_2 a_{21} A_h g_h^n - \tau^2 b_3 a_{31} A_h g_h^n - \tau^2 b_3 a_{32} A_h g_h^{n2} + \frac{1}{6}\tau^3 A_h^2 g_h^n. \end{aligned} \tag{4.15}$$

Proof: This follows by straightforward elimination of the inner stages and by using the simplifying assumption (4.7) and the order conditions (4.9) and (4.10). □

We observe that in the absence of a source term all s-stage methods produce the same approximation, namely

$$u_h^{n+1} = P_s(-\tau A_h)u_h^n, \qquad s = 1, 2, 3, \tag{4.16}$$

where P_s is the so-called *stability polynomial* given by

$$P_s(z) := 1 + z + \ldots \frac{1}{s!}z^s. \tag{4.17}$$

Clearly, in the case without source term the solution to (4.1) is given by

$$u_h(t_{n+1}) = e^{-\tau A_h}u_h(t_n), \tag{4.18}$$

and the RK solution (4.16) can be seen as the approximation of (4.18) obtainded by replacing the exponential function by its $(s, 0)$-Padé approximation.

Furhtermore, we have proven in Section 3.5 that the semi-discrete solution u_h of (4.1) is conservative or dissipative depending on the choice of the fluxes,

$$\|u_h^{\mathrm{cf}}(t)\|_V = 0, \qquad \|u_h^{\mathrm{upw}}(t)\|_V \leq 0,$$

which are properties we would like to mimic in the time discrete case. Obviously, there holds

$$\|u_h^{n+1}\|_V \leq \|P_s(-\tau A_h)\|_{V_h \leftarrow V_h}\|u_h^n\|_V \leq \|P_s(-\tau A_h)\|^n_{V_h \leftarrow V_h}\|u_h^0\|_V,$$

indicating that the stability polynomial is essential for investigating the stability properties of the RK approximation. Clearly, the condition $\|P_s(-\tau A_h)\|_{V_h \leftarrow V_h} \leq 1$ suffices to guarantee stability. One might therefore be attempted to consider the classical spectral stability, i. e. demanding that the eigenvalues of $-\tau A_h$ lie in the stability region $\mathcal{S}$ given by

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

But, it is pointed out in [13] that this stability analysis can be misleading for ill-conditioned problems such as we are dealing with. In particular, the upwind fluxes case cannot be covered by the spectral analysis since it can be shown that the operator A_h^{upw} is non-normal. In contrary, the case of centered fluxes seems to be accessible to this analysis since the operator A_h^{cf} clearly is normal. However, the energy method applies in both cases yielding a full stability analysis and we therefore rely on this method. The stability regions for RK1, RK2 and RK3 are given in Figure 4.1.

4.3 Energy Identities

4.3.1 Homogeneous Energy Identities

In this section we assume that the source term in (4.1) is zero. Then, we have following result for the forward Euler method.

Forward Euler method

Lemma 4.7 *(Homogeneous energy identity for RK1). Let $\{u_h^n\}_n$ be the forward Euler approximation of (4.1). Then, there holds*

$$\|u_h^{n+1}\|_V^2 + 2\tau(u_h^n, A_h u_h^n)_V = \|u_h^n\|_V^2 + \|\tau A_h u_h^n\|_V^2. \tag{4.19}$$

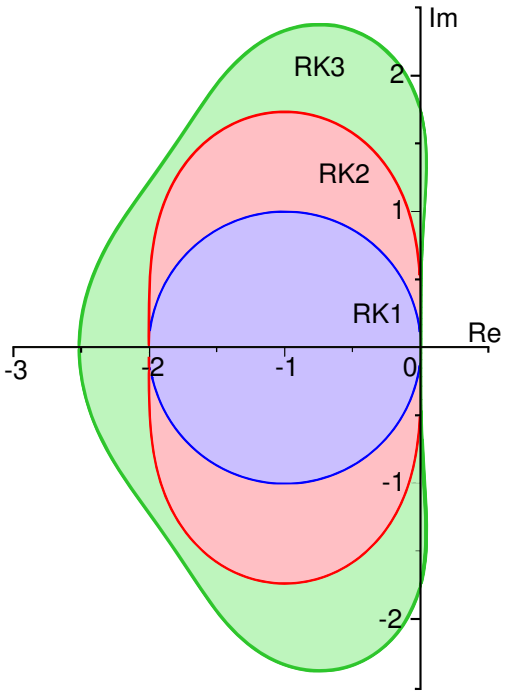


Figure 4.1: Stability regions

Remark 4.8 We see by the appearance of the term $\|\tau A_h u_h^n\|_V^2$ on the RHS of (4.19) that the forward Euler method is *anti-dissipative*, i. e. it produces energy at each time step. In the upwind fluxes case this can be compensated to some extent by the dissipative property (3.65) of the operator A_h^{upw}. In fact, in the upwind fluxes case we have the following energy identity,

$$\|u_h^{n+1}\|_V^2 + 2\tau|u_h^n|_S^2 = \|u_h^n\|_V^2 + \|\tau A_h^{\mathrm{upw}} u_h^n\|_V^2.$$

In contrary, owing to the skew-symmetry of A_h^{cf}, see (3.64), the energy identity in the centered fluxes case reads

$$\|u_h^{n+1}\|_V^2 = \|u_h^n\|_V^2 + \|\tau A_h^{\mathrm{cf}} u_h^n\|_V^2.$$

◇

Proof: We calculate the norm of u_h^{n+1} using the recursion (4.13),

$$\|u_h^{n+1}\|_V^2 = \|u_h^n - \tau A_h u_h^n\|_V^2 = \|u_h^n\|_V^2 - 2(u_h^n, \tau A_h u_h^n)_V + \|\tau A_h u_h^n\|_V^2.$$

□

RK2 methods We proceed with RK2 methods. We first observe that we can write the RK2 approximation as a corrected forward Euler step. Let therefore $\widetilde{U}_h^{n1}$ denote the forward Euler step, i. e.

$$\widetilde{U}_h^{n1} := u_h^n - \tau A_h u_h^n. \tag{4.20}$$

Then, the RK2 approximation (4.14) can be written as

$$u_h^{n+1} = \widetilde{U}_h^{n1} + \frac{1}{2}\tau^2 A_h^2 u_h^n = \widetilde{U}_h^{n1} + \frac{1}{2}\tau A_h(u_h^n - \widetilde{U}_h^{n1}). \tag{4.21}$$

We use this identity to state the RK2 energy identity.

Lemma 4.9 *(Homogeneous energy identity for RK2). Let* $\{u_h^n\}_n$ *be a RK2 approximation to (4.1). Then, there holds*

$$\|u_h^{n+1}\|_V^2 + \tau(u_h^n, A_h u_h^n)_V + \tau(\widetilde{U}_h^{n1}, A_h \widetilde{U}_h^{n1})_V = \|u_h^n\|_V^2 + \frac{1}{4}\|\tau^2 A_h^2 u_h^n\|_V^2. \tag{4.22}$$

Remark 4.10 Again, we observe an anti-dissipative behaviour but this time induced by the term $\frac{1}{4}\|\tau^2 A_h^2 u_h^n\|_V^2$. Furthermore, comparing (4.22) with (4.19) we see that in the upwind fluxes case the RK2 approximation additionaly yields the dissipative term $|\widetilde{U}_h^{n1}|_S^2$. Indeed, there holds

$$\|u_h^{n+1}\|_V^2 + \tau|u_h^n|_S^2 + \tau|\widetilde{U}_h^{n1}|_S^2 = \|u_h^n\|_V^2 + \frac{1}{4}\|\tau^2 (A_h^{\mathrm{upw}})^2 u_h^n\|_V^2.$$

In the case of centered fluxes we have

$$\|u_h^{n+1}\|_V^2 = \|u_h^n\|_V^2 + \frac{1}{4}\|\tau^2 (A_h^{\mathrm{cf}})^2 u_h^n\|_V^2.$$

◇

Proof: We compute the norm of u_h^{n+1} using the recursion (4.21),

$$\|u_h^{n+1}\|_V^2 = \|\widetilde{U}_h^{n1}\|_V^2 + (\widetilde{U}_h^{n1}, \tau A_h u_h^n)_V - (\widetilde{U}_h^{n1}, \tau A_h \widetilde{U}_h^{n1})_V + \frac{1}{4}\|\tau A_h (u_h^n - \widetilde{U}_h^{n1})\|_V^2.$$

Applying the RK1 energy identity (4.19) to $\|\widetilde{U}_h^{n1}\|_V^2$ yields

$$\begin{aligned}\|u_h^{n+1}\|_V^2 &+ (u_h^n, \tau A_h u_h^n)_V + (\widetilde{U}_h^{n1}, \tau A_h \widetilde{U}_h^{n1})_V \\ &= \|u_h^n\|_V^2 + (\widetilde{U}_h^{n1} - u_h^n, \tau A_h u_h^n)_V + \|\tau A_h u_h^n\|_V^2 + \frac{1}{4}\|\tau A_h (u_h^n - \widetilde{U}_h^{n1})\|_V^2.\end{aligned}$$

Using (4.20) we see

$$(\widetilde{U}_h^{n1} - u_h^n, \tau A_h u_h^n)_V = -\|\tau A_h u_h^n\|_V^2,$$

and

$$\frac{1}{4}\|\tau A_h (u_h^n - \widetilde{U}_h^{n1})\|_V^2 = \frac{1}{4}\|\tau^2 A_h^2 u_h^n\|_V^2,$$

which proves the claim. □

RK3 methods For RK3 methods we define $\widetilde{U}_h^{n2}$ as the approximation given by a RK2 scheme, i. e.

$$\widetilde{U}_h^{n2} := u_h^n - \tau A_h u_h^n + \frac{1}{2}\tau^2 A_h^2 u_h^n = \widetilde{U}_h^{n1} + \frac{1}{2}\tau^2 A_h^2 u_h^n. \tag{4.23}$$

Revisiting the recursion for RK3 methods (4.15) we see that we can write the RK3 approximation as

$$u_h^{n+1} = \widetilde{U}_h^{n2} - \frac{1}{6}\tau^3 A_h^3 u_h^n = \widetilde{U}_h^{n2} + \frac{1}{3}\tau A_h (\widetilde{U}_h^{n1} - \widetilde{U}_h^{n2}). \tag{4.24}$$

Lemma 4.11 *(Homogeneous energy identity for RK3). Let* $\{u_h^n\}_n$ *be a RK3 approximation to (4.1). Then, there holds*

$$\begin{aligned}\|u_h^{n+1}\|_V^2 + \tau(u_h^n, A_h u_h^n)_V + \frac{1}{3}\tau(\widetilde{U}_h^{n1}, A_h \widetilde{U}_h^{n1})_V + \frac{2}{3}\tau(\widetilde{U}_h^{n2}, A_h \widetilde{U}_h^{n2})_V + \frac{1}{12}\|\tau^2 A_h^2 u_h^n\|_V^2 \\ = \|u_h^n\|_V^2 + \frac{1}{3}\tau(\tau A_h u_h^n, \tau A_h^2 u_h^n)_V + \frac{1}{36}\|\tau^3 A_h^3 u_h^n\|_V^2.\end{aligned} \tag{4.25}$$

Remark 4.12 We see that RK3 schemes essentially differ from RK1 and RK2 schemes. Even in the centered fluxes case we have a dissipative term. Indeed, there holds

$$\|u_h^{n+1}\|_V^2 + \frac{1}{12}\|\tau^2 (A_h^{\mathrm{cf}})^2 u_h^n\|_V^2 = \|u_h^n\|_V^2 + \frac{1}{36}\|\tau^3 (A_h^{\mathrm{cf}})^3 u_h^n\|_V^2.$$

We will see later that this significantly improves the stability behaviour. For the upwind fluxes case the energy identity (4.25) reads as

$$\begin{aligned}\|u_h^{n+1}\|_V^2 + \tau|u_h^n|_S^2 + \frac{1}{3}\tau|\widetilde{U}_h^{n1}|_S^2 + \frac{2}{3}\tau|\widetilde{U}_h^{n2}|_S^2 + \frac{1}{12}\|\tau^2 (A_h^{\mathrm{upw}})^2 u_h^n\|_V^2 \\ = \|u_h^n\|_V^2 + \frac{1}{3}\tau|\tau A_h^{\mathrm{upw}} u_h^n|_S^2 + \frac{1}{36}\|\tau^3 (A_h^{\mathrm{upw}})^3 u_h^n\|_V^2.\end{aligned}$$

Compared to the centered fluxes case we benefit on the one hand from the additional dissipation provided from space discretization in form of the S-seminorm terms on the LHS but on the other hand suffer from the non-negative term $\frac{1}{3}\tau|\tau A_h u_h^n|_S^2$ on the RHS. ◇

Proof: We use the recursion (4.24) to infer

$$\|u_h^{n+1}\|_V^2 = \|\widetilde{U}_h^{n2}\|_V^2 + \frac{2}{3}(\widetilde{U}_h^{n2}, \tau A_h \widetilde{U}_h^{n1})_V - \frac{2}{3}(\widetilde{U}_h^{n2}, \tau A_h \widetilde{U}_h^{n2})_V + \frac{1}{9}\|\tau A_h(\widetilde{U}_h^{n1} - \widetilde{U}_h^{n2})\|_V^2.$$

Applying the RK2 energy identity (4.22) to the $\|\widetilde{U}_h^{n2}\|_V^2$ we get

$$\begin{aligned}\|u_h^{n+1}\|_V^2 + (u_h^n, \tau A_h u_h^n)_V + \frac{1}{3}(\widetilde{U}_h^{n1}, \tau A_h \widetilde{U}_h^{n1})_V + \frac{2}{3}(\widetilde{U}_h^{n2}, \tau A_h \widetilde{U}_h^{n2})_V \\ = \|u_h^n\|_V^2 + \frac{2}{3}(\widetilde{U}_h^{n2} - \widetilde{U}_h^{n1}, \tau A_h \widetilde{U}_h^{n1})_V + \frac{1}{4}\|\tau^2 A_h u_h^n\|_V^2 + \frac{1}{9}\|\tau A_h(\widetilde{U}_h^{n1} - \widetilde{U}_h^{n2})\|_V^2. \quad (4.26)\end{aligned}$$

Let us consider the second term on the RHS. From (4.23) we have $\widetilde{U}_h^{n2} - \widetilde{U}_h^{n1} = \frac{1}{2}\tau^2 A_h^2 u_h^n$ and thus we deduce

$$\frac{2}{3}(\widetilde{U}_h^{n2} - \widetilde{U}_h^{n1}, \tau A_h \widetilde{U}_h^{n1})_V = \frac{2}{3}(\tau^2 A_h^2 u_h^n, \tau A_h u_h^n)_V - \frac{1}{3}\|\tau^2 A_h^2 u_h^n\|_V^2,$$

as well as

$$\frac{1}{9}\|\tau A_h(\widetilde{U}_h^{n1} - \widetilde{U}_h^{n2})\|_V^2 = \frac{1}{36}\|\tau^3 A_h^3 u_h^n\|_V^2.$$

Inserting this equalities in (4.26) yields (4.25). □

4.3.2 Inhomogeneous Energy Identities

We extend the ideas from the previous section to the general case with $g_h \neq 0$. In the case of the forward Euler method this can be done straightforward.

Forward Euler method

Lemma 4.13 *(Energy identity for RK1). Let $\{u_h^n\}_n$ be the forward Euler discretization of (4.1). Then, there holds*

$$\|u_h^{n+1}\|_V^2 + 2\tau(u_h^n, A_h u_h^n)_V = \|u_h^n\|_V^2 + 2\tau(u_h^n, g_h^n)_V + \|\tau A_h u_h^n - \tau g_h^n\|_V^2. \quad (4.27)$$

Proof: Using the recursion (4.13) we see

$$\|u_h^{n+1}\|_V^2 = \|u_h^n - \tau A_h u_h^n + \tau g_h^n\|_V^2 = \|u_h^n\| - 2(u_h^n, \tau A_h u_h^n - \tau g_h^n)_V + \|\tau A_h u_h^n - \tau g_h^n\|_V^2.$$

□

RK2 methods Reviewing the homogeneous RK2 energy identity (4.22) we realize that it was important to write the RK2 scheme as a corrected forward Euler method. From the RK2 recursion (4.14) it is not clear if this is possible if the source term is non-zero. We show that this can indeed be done. First, we introduce Peano kernels, which provide a uniform notation of the remainder terms of Taylor expansions.

Definition 4.14 (Peano kernel). We define the Peano kernel with parameters x and p as

$$\kappa_{x,p}(t) := \frac{1}{p!}(x-t)_+^p = \begin{cases} \frac{1}{p!}(x-t)^p, & t \le x, \\ 0, & t > x. \end{cases} \tag{4.28}$$

Now, let U_h^{n1} denote the forward Euler step, i. e.

$$U_h^{n1} := u_h^n - \tau A_h u_h^n + \tau g_h^n. \tag{4.29}$$

Then, we can prove the following RK2 recursion.

Lemma 4.15 *(RK2 recursion). Let $g_h \in C^2(0,T;V_h)$. Then, we can write the RK2 approximation (4.14) as*

$$u_h^{n+1} = U_h^{n1} + \frac{1}{2}\tau A_h(u_h^n - U_h^{n1}) + \frac{1}{2}\tau^2 \partial_t g_h^n + \tau R_2^n, \tag{4.30}$$

where the remainder term R_2^n is given as

$$R_2^n = -\tau b_2 \int_{t_n}^{t_{n+1}} \kappa'_{c_2,2}\left(\frac{s-t_n}{\tau}\right) \partial_{tt} g_h(s)\, ds.$$

Proof: Recall that we have shown in Lemma 4.6 the RK2 recursion

$$u_h^{n+1} = u_h^n - \tau A_h u_h^n + \frac{1}{2}\tau^2 A_h^2 u_h^n + \tau b_1 g_h^n + \tau b_2 g_h^{n2} - \frac{1}{2}\tau^2 A_h g_h^n. \tag{4.31}$$

Employing Taylor expansion to g_h^{n2} yields

$$g_h^{n2} = g_h^n + c_2\tau \partial_t g_h^n - \tau \int_{t_n}^{t_{n+1}} \kappa'_{c_2,2}\left(\frac{s-t_n}{\tau}\right) \partial_{tt} g_h(s)\, ds.$$

Inserting this in (4.31) gives

$$u_h^{n+1} = u_h^n - \tau A_h u_h^n + \frac{1}{2}\tau^2 A_h^2 u_h^n + \tau(b_1+b_2) g_h^n + \tau^2 b_2 c_2 \partial_t g_h^n - \frac{1}{2}\tau^2 A_h g_h^n + \tau R_2^n.$$

Using the simplifying assumption (4.7) and the RK2 order condition (4.9) it follows

$$u_h^{n+1} = u_h^n - \tau A_h u_h^n + \tau g_h^n + \frac{1}{2}\tau A_h(\tau A_h u_h^n - \tau g_h^n) + \frac{1}{2}\tau^2 \partial_t g_h^n + \tau R_2^n,$$

whence we infer (4.30) by (4.29). □

This lemma enables us to prove the RK2 energy identity for the inhomogeneous case.

Lemma 4.16 *(Energy identity for RK2). Let $\{u_h^n\}_n$ be a RK2 approximation of (4.1). Then, there holds*

$$\begin{aligned} &\|u_h^{n+1}\|_V^2 + \tau(u_h^n, A_h u_h^n)_V + \tau(U_h^{n1}, A_h U_h^{n1})_V \\ &\quad = \|u_h^n\|_V^2 + \tau(u_h^n, g_h^n)_V + \tau(U_h^{n1}, g_h^n + \tau \partial_t g_h^n + 2R_2^n)_V \\ &\quad\quad + \frac{1}{4}\|\tau^2 A_h^2 u_h^n - \tau^2 A_h g_h^n + \tau^2 \partial_t g_h^n + 2\tau R_2^n\|_V^2 \end{aligned} \tag{4.32}$$

Proof: We abbreviate the last two terms in the recursion (4.30) with R, i. e.

$$u_h^{n+1} = U_h^{n1} + \frac{1}{2}\tau A_h(u_h^n - U_h^{n1}) + R.$$

Taking the inner-product of the above equation with itself we get

$$\begin{aligned}\|u_h^{n+1}\|_V^2 =& \|U_h^{n1}\|_V^2 + (U_h^{n1}, \tau A_h u_h^n)_V - (U_h^{n1}, \tau A_h U_h^{n1})_V + 2(U_h^{n1}, R)_V \\ &+ \frac{1}{4}\|\tau A_h(u_h^n - U_h^{n1}) + 2R\|_V^2.\end{aligned}$$

Using the RK1 identity (4.27) for the term $\|U_h^{n1}\|_V^2$ yields

$$\begin{aligned}\|u_h^{n+1}\|_V^2 &+ (u_h^n, \tau A_h u_h^n)_V + (U_h^{n1}, \tau A_h U_h^{n1})_V \\ &= \|u_h^n\|_V^2 + 2(u_h^n, \tau g_h^n)_V + 2(U_h^{n1}, R)_V + (U_h^{n1} - u_h^n, \tau A_h u_h^n)_V \\ &\quad + \|\tau A_h u_h^n - \tau g_h^n\|_V^2 + \frac{1}{4}\|\tau A_h(u_h^n - U_h^{n1}) + 2R\|_V^2.\end{aligned}$$

By (4.29) we see $U_h^{n1} - u_h^n = -\tau A_h u_h^n + \tau g_h^n$ and thus that it holds

$$(U_h^{n1} - u_h^n, \tau A_h u_h^n)_V = -\|\tau A_h u_h^n - \tau g_h^n\|_V^2 + (U_h^{n1} - u_h^n, \tau g_h^n)_V,$$

whence we conclude (4.32). □

RK3 methods We proceed with RK3 schemes by the same steps as for RK2 methods. First we introduce

$$U_h^{n2} := U_h^{n1} + \frac{1}{2}\tau A_h(u_h^n - U_h^{n1}) + \frac{1}{2}\tau^2 \partial_t g_h^n, \tag{4.33}$$

which is the RK2 step without the remainder term R_2^n. This allows to prove the following RK3 recursion.

Lemma 4.17 *(RK3 recursion). Let $g_h \in C^3(0,T;V_h)$. Then, we can write the RK3 approximation (4.15) as*

$$u_h^{n+1} = U_h^{n2} + \frac{1}{3}\tau A_h(U_h^{n1} - U_h^{n2}) + \frac{1}{6}\tau^3 \partial_{tt} g_h^n + \tau R_3^n, \tag{4.34}$$

where the remainder term R_3^n is given as

$$R_3^n := -\tau^2 \sum_{i=2}^{3} \left[b_i \int_{t_n}^{t_{n+1}} \kappa'_{c_i,3}\left(\frac{s-t_n}{\tau}\right) \partial_{ttt} g_h(s)\, ds \right] - \tau^2 b_3 a_{32} \int_{t_n}^{t_{n+1}} \kappa''_{c_2,3}\left(\frac{s-t_n}{\tau}\right) A_h(\partial_{tt} g_h(s))\, ds.$$

Note that for RK3 methods the remainder term includes the term $A_h(\partial_{tt} g_h)$. This will later force us to demand higher regularity assumptions than $g_h \in C^3(0,T;V_h)$.

Proof: Recall that the RK3 recursion (4.15) reads

$$\begin{aligned}u_h^{n+1} =& u_h^n - \tau A_h u_h^n + \frac{1}{2}\tau^2 A_h^2 u_h^n - \frac{1}{6}\tau^3 A_h^3 u_h^n + \tau b_1 g_h^n + \tau b_2 g_h^{n2} + \tau b_3 g_h^{n3} \\ &- \tau^2 b_2 a_{21} A_h g_h^n - \tau^2 b_3 a_{31} A_h g_h^n - \tau^2 b_3 a_{32} A_h g_h^{n2} + \frac{1}{6}\tau^3 A_h^2 g_h^n.\end{aligned} \tag{4.35}$$

We expand g_h^{2n} and g_h^{3n} in the above equation into a second order Taylor series,

$$g_h^{ni} = g_h^n + \tau c_i \partial_t g_h^n + \frac{1}{2}\tau^2 c_i^2 \partial_{tt} g_h^n - \tau^2 \int_{t_n}^{t_{n+1}} \kappa'_{c_i,3}\left(\frac{s-t_n}{\tau}\right) \partial_{ttt} g_h(s)\, ds, \qquad i = 2,3,$$

and g_h^{2n} stemming from $A_h g_h^{n2}$ into a first order Taylor series,

$$g_h^{2n} = g_h^n + \tau c_2 \partial_t g_h^n + \tau \int_{t_n}^{t_{n+1}} \kappa''_{c_2,3}\left(\frac{s-t_n}{\tau}\right) \partial_{tt} g_h(s)\, ds.$$

Then, the source terms in (4.35) can be written as

$$\begin{aligned}&\tau(b_1+b_2+b_3)g_h^n + \tau^2(b_2c_2+b_3c_3)\partial_t g_h^n + \frac{1}{2}\tau^3(b_2c_2^2+b_3c_3^2)\partial_{tt} g_h^n \\ &- \tau^2(b_2a_{21}+b_3a_{31}+b_3a_{32})A_h g_h^n - \tau^3 b_3 a_{32} c_2 A_h(\partial_t g_h^n) + \frac{1}{6}\tau^3 A_h^2 g_h^n + \tau R_3^n. \end{aligned} \tag{4.36}$$

Inserting (4.36) together with the simplifying assumption (4.7) and the RK3 order conditions (4.10) into (4.35) gives

$$\begin{aligned}u_h^{n+1} =& u_h^n - \tau A_h u_h^n + \frac{1}{2}\tau^2 A_h^2 u_h^n - \frac{1}{6}\tau^3 A_h^3 u_h^n + \tau g_h^n + \frac{1}{2}\tau^2 \partial_t g_h^n + \frac{1}{6}\tau^3 \partial_{tt} g_h^n \\ &- \frac{1}{2}\tau^2 A_h g_h^n - \frac{1}{6}\tau^3 A_h(\partial_t g_h^n) + \frac{1}{6}\tau^3 A_h^2 g_h^n + \tau R_3^n.\end{aligned}$$

Inserting (4.29) in (4.33) yields

$$U_h^{n2} = u_h^n - \tau A_h u_h^n + \frac{1}{2}\tau^2 A_h^2 u_h^n + \tau g_h^n + \frac{1}{2}\tau^2 \partial_t g_h^n - \frac{1}{2}\tau^2 A_h g_h^n,$$

whence we deduce

$$u_h^{n+1} = U_h^{n2} - \frac{1}{6}\tau^3 A_h^3 u_h^n - \frac{1}{6}\tau^3 A_h(\partial_t g_h^n) + \frac{1}{6}\tau^3 A_h^2 g_h^n + \frac{1}{6}\tau^3 \partial_{tt} g_h^n + \tau R_3^n.$$

We can write this equation as

$$\begin{aligned}u_h^{n+1} =& U_h^{n2} - \frac{1}{3}\tau A_h\left(u_h^n - \tau A_h u_h^n + \frac{1}{2}A_h^2 u_h^n + \tau g_h^n + \frac{1}{2}\tau^2 \partial_t g_h^n - \frac{1}{2}\tau^2 A_h g_h^n\right) \\ &+ \frac{1}{3}\tau A_h\left(u_h^n - \tau A_h u_h^n + \tau g_h^n\right) + \frac{1}{6}\tau^3 \partial_{tt} g_h^n + \tau R_3^n \\ =& U_h^{n2} - \frac{1}{3}\tau A_h U_h^{n2} + \frac{1}{3}\tau A_h U_h^{n1} + \frac{1}{6}\tau^3 \partial_{tt} g_h^n + \tau R_3^n.\end{aligned}$$

□

We end this section with the energy identity for RK3 schemes.

Lemma 4.18 *(Energy identity for RK3). Let $\{u_h^n\}_n$ be a RK3 approximation of (4.1). Then, there holds*

$$\begin{aligned}\|u_h^{n+1}\|_V^2 +& \tau(u_h^n, A_h u_h^n)_V + \frac{1}{3}\tau(U_h^{n1}, A_h U_h^{n1})_V \\ &+ \frac{2}{3}\tau(U_h^{n2}, A_h U_h^{n2})_V + \frac{1}{12}\|\tau^2 A_h^2 u_h^n - \tau^2 A_h g_h^n + \tau^2 \partial_t g_h^n\|_V^2 \\ =& \|u_h^n\|_V^2 + \frac{1}{3}\tau\big((\tau A_h u_h^n - \tau g_h^n), A_h(\tau A_h u_h^n - \tau g_h^n)\big)_V \\ &+ \tau(u_h^n, g_h^n + \frac{1}{3}\tau\partial_t g_h^n)_V + \frac{1}{3}\tau(U_h^{n1}, g_h^n)_V + \frac{2}{3}\tau(U_h^{n2}, g_h^n + \tau\partial_t g_h^n + \frac{1}{2}\tau^2\partial_{tt} g_h^n + 3R_3^n)_V \\ &+ \frac{1}{36}\|\tau^3 A_h^3 u_h^n - \tau^3 A_h^2 g_h^n + \tau^3 A_h(\partial_t g_h^n) - \tau^3 \partial_{tt} g_h^n - 6\tau R_3^n\|_V^2.\end{aligned} \tag{4.37}$$

Proof: By denoting $R := \frac{1}{2}\tau^3\partial_{tt}g_h^n + 3\tau R_3^n$ we can write the RK3 recursion (4.34) as

$$u_h^{n+1} = U_h^{n2} + \frac{1}{3}\tau A_h(U_h^{n1} - U_h^{n2}) + \frac{1}{3}R.$$

By taking the inner-product of the above equation with itself we get

$$\begin{aligned}\|u_h^{n+1}\|_V^2 =& \|U_h^{n2}\|_V^2 + \frac{2}{3}(U_h^{n2}, \tau A_h U_h^{n1})_V - \frac{2}{3}(U_h^{n2}, \tau A_h U_h^{n2})_V + \frac{2}{3}(U_h^{n2}, R)_V \\ &+ \frac{1}{9}\|\tau A_h(U_h^{n1} - U_h^{n2}) + R\|_V^2. \end{aligned} \tag{4.38}$$

We can use the RK2 energy identity (4.32) for $\|U_h^{n2}\|_V^2$ by dropping the remainder term R_2^n. This yields

$$\begin{aligned}\|U_h^{n2}\|_V^2 + \tau(u_h^n, A_h u_h^n)_V + \tau(U_h^{n1}, A_h U_h^{n1})_V =& \|u_h^n\|_V^2 + \tau(u_h^n, g_h^n)_V + \tau(U_h^{n1}, g_h^n + \tau\partial_t g_h^n)_V \\ &+ \frac{1}{4}\|\tau^2 A_h^2 u_h^n - \tau^2 A_h g_h^n + \tau^2\partial_t g_h^n\|_V^2.\end{aligned}$$

Plugging this into (4.38) gives

$$\begin{aligned}\|u_h^{n+1}\|_V^2 +& (u_h^n, \tau A_h u_h^n)_V + \frac{1}{3}(U_h^{n1}, \tau A_h U_h^{n1})_V + \frac{2}{3}(U_h^{n2}, \tau A_h U_h^{n2})_V \\ =& \|u_h^n\|_V^2 + (u_h^n, \tau g_h^n)_V + (U_h^{n1}, \tau g_h^n + \tau^2\partial_t g_h^n)_V + \frac{2}{3}(U_h^{n2}, R)_V \\ &+ \frac{2}{3}(U_h^{n2} - U_h^{n1}, \tau A_h U_h^{n1})_V + \frac{1}{4}\|\tau^2 A_h^2 u_h^n - \tau^2 A_h g_h^n + \tau^2\partial_t g_h^n\|_V^2 \\ &+ \frac{1}{9}\|\tau A_h(U_h^{n1} - U_h^{n2}) + R\|_V^2. \end{aligned} \tag{4.39}$$

Let us consider the term $\frac{2}{3}(U_h^{n2} - U_h^{n1}, \tau A_h U_h^{n1})_V$. It holds

$$U_h^{n2} - U_h^{n1} = \frac{1}{2}\tau^2 A_h^2 u_h^n - \frac{1}{2}\tau^2 A_h g_h^n + \frac{1}{2}\tau^2\partial_t g_h^n, \tag{4.40}$$

and

$$\tau A_h U_h^{n1} = \tau A_h(u_h^n - \tau A_h u_h^n + \tau g_h^n) = (\tau A_h u_h^n + \tau^2\partial_t g_h^n) - (\tau^2 A_h^2 u_h^n - \tau^2 A_h g_h^n + \tau^2\partial_t g_h^n).$$

Consequently we have

$$\begin{aligned}\frac{2}{3}(U_h^{n2} - U_h^{n1}, \tau A_h U_h^{n1})_V =& \frac{1}{3}(\tau^2 A_h^2 u_h^n - \tau^2 A_h g_h^n + \tau^2\partial_t g_h^n, \tau A_h u_h^n + \tau^2\partial_t g_h^n)_V \\ &- \frac{1}{3}\|\tau^2 A_h^2 u_h^n - \tau^2 A_h g_h^n + \tau^2\partial_t g_h^n\|_V^2. \end{aligned} \tag{4.41}$$

The first term is yet unpleasant and we rewrite it as

$$\begin{aligned}\frac{1}{3}(\tau^2 A_h^2 u_h^n -& \tau^2 A_h g_h^n + \tau^2\partial_t g_h^n, \tau A_h u_h^n + \tau^2\partial_t g_h^n)_V \\ =& \frac{1}{3}\big(\tau A_h(\tau A_h u_h^n - \tau g_h^n), (\tau A_h u_h^n - \tau g_h^n)\big)_V + \frac{1}{3}\big(\tau^2\partial_t g_h^n, (\tau A_h u_h^n - \tau g_h^n)\big)_V \\ &+ \frac{1}{3}(\tau^2 A_h^2 u_h^n - \tau^2 A_h g_h^n + \tau^2\partial_t g_h^n, \tau g_h^n + \tau^2\partial_t g_h^n)_V, \end{aligned} \tag{4.42}$$

which ensure that we can at least apply the skew-adjointness of A_h^{cf} or the dissipative property of A_h^{upw} to the first term. Furthermore, from $u_h^n - U_h^{n1} = \tau A_h u_h^n - \tau g_h^n$ and (4.40) we see that the last two terms in (4.42) can be written as

$$\begin{aligned}\frac{1}{3}\big(\tau^2\partial_t g_h^n, (\tau A_h u_h^n - \tau g_h^n)\big)_V &+ \frac{1}{3}(\tau^2 A_h^2 u_h^n - \tau^2 A_h g_h^n + \tau^2\partial_t g_h^n, \tau g_h^n + \tau^2\partial_t g_h^n)_V \\ =& \frac{1}{3}(u_h^n, \tau^2\partial_t g_h^n)_V - (U_h^{n1}, \frac{2}{3}g_h^n + \tau^2\partial_t g_h^n)_V + \frac{2}{3}(U_h^{n2}, \tau g_h^n + \tau^2\partial_t g_h^n)_V. \end{aligned} \tag{4.43}$$

Inserting (4.41) with (4.42) and (4.43) in (4.39) yields

$$\begin{aligned}\|u_h^{n+1}\|_V^2 &+ (u_h^n, \tau A_h u_h^n)_V + \frac{1}{3}(U_h^{n1}, \tau A_h U_h^{n1})_V + \frac{2}{3}(U_h^{n2}, \tau A_h U_h^{n2})_V \\ &= \|u_h^n\|_V^2 + (u_h^n, \tau g_h^n + \frac{1}{3}\tau^2 \partial_t g_h^n)_V + \frac{1}{3}(U_h^{n1}, \tau g_h^n)_V + \frac{2}{3}(U_h^{n2}, \tau g_h^n + \tau^2 \partial_t g_h^n + R)_V \\ &+ \big(\tau A_h(\tau A_h u_h^n - \tau g_h^n), (\tau A_h u_h^n - \tau g_h^n)\big)_V - \frac{1}{12}\|\tau^2 A_h^2 u_h^n - \tau^2 A_h g_h^n + \tau^2 \partial_t g_h^n\|_V^2 \\ &+ \frac{1}{9}\|\tau A_h(U_h^{n1} - U_h^{n2}) + \tau R\|_V^2.\end{aligned}$$

Employing the identity (4.40) in the last term concludes the proof. □

4.4 Stability

Using the energy identities we can prove the stability of the RK discretizations. We will see that each method is stable only if the time step size is bounded w. r. t. the meshsize h. This condition is called the *CFL condition*. Furthermore, we will see that higher order methods admit better stability properties in the sense of more relaxed CFL conditions. From now on, we assume without loss of generality that $\tau \leq 1$.

Definition 4.19 (CFL-conditions). Let ϱ, ϱ' and ϱ'' be positive numbers. We say that the step size τ satisfies the *usual CFL condition* if it holds

$$\tau \leq \varrho \frac{h}{c_\infty}. \tag{4.44}$$

Furthermore, the step size satisfies the *4/3-CFL condition* if

$$\tau \leq \varrho' \left(\frac{h}{c_\infty}\right)^{4/3}, \tag{4.45}$$

and it satisfies the *2-CFL condition* if

$$\tau \leq \varrho'' \left(\frac{h}{c_\infty}\right)^{2}. \tag{4.46}$$

The hierachy of the CFL conditions is as follows: The 2-CFL is the strongest assumption and implies the 4/3-CFL condition and the usual CFL condition. Furthermore, the 4/3-CFL condition implies the usual CFL condition.

Our first stability result is for the forward Euler scheme which requires the strongest 2-CFL condition.

Forward Euler method

Lemma 4.20 *(Stability for RK1). Let $\{u_h^n\}_n$ be the forward Euler approximation of (4.1). Then, under the 2-CFL condition (4.46), the following results hold:*

i) In the case of centered fluxes,

$$\|u_h^n\|_V^2 \leq C_1 \left(\|u_h^0\|_V^2 + 3\tau \sum_{m=0}^{n-1} \|g_h^m\|_V^2\right). \tag{4.47}$$

ii) In the case of upwind fluxes,

$$\|u_h^n\|_V^2 + 2\tau \sum_{m=0}^{n-1} |u_h^m|_S^2 \leq C_1' \left(\|u_h^0\|_V^2 + 3\tau \sum_{m=0}^{n-1} \|g_h^m\|_V^2\right). \tag{4.48}$$

The constants are $C_1 = \exp\big((1 + 2(C_h^{\mathrm{cf}})^2 \varrho'')t_n\big)$ and $C_1' = \exp\big((1 + 2(C_h^{\mathrm{upw}})^2 \varrho'')t_n\big)$.

Remark 4.21 Note that $t_n \leq T$ guarantees that the constants C_1 and C_1' are uniformly bounded w. r. t. τ (and clearly also w. r. t. h). Furthermore, recall that by the boundedness of the projection operator (3.59) we have $\|u_h^0\|_V \leq \|u_0\|_V$ and $\|g_h^m\|_V \leq \|g(t_m)\|_V$. In summary, the inequalities (4.47) and (4.48) guarantee the continuous dependency of the fully discrete solution on the initial value u_0 and the source term g independent of both h and τ. However, the stability constants depend exponentially on the 2-CFL condition constant ϱ'' and on the constants C_h^{cf} and C_h^{upw}. ◇

Proof: We recall the energy identity (4.27) for the forward Euler method:

$$\|u_h^{n+1}\|_V^2 + 2\tau(u_h^n, A_h u_h^n)_V = \|u_h^n\|_V^2 + 2\tau(u_h^n, g_h^n)_V + \|\tau A_h u_h^n - \tau g_h^n\|_V^2.$$

We apply the Cauchy-Schwarz inequality and Young's inequality to the second term of the RHS and the triangle inequality and Young's inequality to the last term,

$$\|u_h^{n+1}\|_V^2 + 2\tau(u_h^n, A_h u_h^n)_V \leq \|u_h^n\|_V^2 + \tau\|u_h^n\|_V^2 + \tau\|g_h^n\|_V^2 + 2\|\tau A_h u_h^n\|_V^2 + 2\|\tau g_h^n\|_V^2.$$

The term $\|\tau A_h u_h^n\|_V^2$ can be estimated with the bound (4.2) of the discrete operator A_h,

$$\|u_h^{n+1}\|_V^2 + 2\tau(u_h^n, A_h u_h^n)_V \leq \|u_h^n\|_V^2 + \tau\|u_h^n\|_V^2 + (1+2\tau)\tau\|g_h^n\|_V^2 + 2C_h^2 c_\infty^2 \tau^2 h^{-2}\|u_h^n\|_V^2.$$

Applying the 2-CFL condition and using $(1+2\tau) \leq 3$ then yields

$$\|u_h^{n+1}\|_V^2 - \|u_h^n\|_V^2 + 2\tau(u_h^n, A_h u_h^n)_V \leq 3\tau\|g_h^n\|_V^2 + (1 + 2C_h^2\varrho'')\tau\|u_h^n\|_V^2. \tag{4.49}$$

Let us denote $C_0 := 1 + 2C_h^2\varrho''$. Summing (4.49) from 0 to $n-1$ gives

$$\|u_h^n\|_V^2 + 2\tau\sum_{m=0}^{n-1}(u_h^m, A_h u_h^m)_V \leq \|u_h^0\|_V^2 + 3\tau\sum_{m=0}^{n-1}\|g_h^m\|_V^2 + C_0\tau\sum_{m=0}^{n-1}\|u_h^m\|_V^2, \tag{4.50}$$

which obviously implies

$$\|u_h^n\|_V^2 \leq \|u_h^0\|_V^2 + 3\tau\sum_{m=0}^{n-1}\|g_h^m\|_V^2 + C_0\tau\sum_{m=0}^{n-1}\|u_h^m\|_V^2.$$

This inequality meets the assumptions of the discrete Gronwall Lemma A.5 and its application provides the estimate

$$\|u_h^n\|_V^2 \leq e^{C_0 n\tau}\left(\|u_h^0\|_V^2 + 3\tau\sum_{m=0}^{n-1}\|g_h^m\|_V^2\right).$$

Inserting this bound in the RHS of (4.50) shows

$$\|u_h^n\|_V^2 + 2\tau\sum_{m=0}^{n-1}(u_h^m, A_h u_h^m)_V \leq \|u_h^0\|_V^2 + 3\tau\sum_{m=0}^{n-1}\|g_h^m\|_V^2 + C_0\tau\sum_{m=0}^{n-1}\left[e^{C_0 m\tau}\left(\|u_h^0\|_V^2 + 3\tau\sum_{l=0}^{m-1}\|g_h^l\|_V^2\right)\right],$$

which can be estimated by

$$\|u_h^n\|_V^2 + 2\tau\sum_{m=0}^{n-1}(u_h^m, A_h u_h^m)_V \leq \left(1 + C_0\tau\sum_{m=0}^{n-1}e^{C_0 m\tau}\right)\left(\|u_h^0\|_V^2 + 3\tau\sum_{m=0}^{n-1}\|g_h^m\|_V^2\right).$$

Note that the sum $\tau\sum_{m=0}^{n-1}e^{C_0 m\tau}$ is a lower sum of the monotonically increasing function $e^{C_0 t}$ and thus we can deduce

$$\tau\sum_{m=0}^{n-1}e^{C_0 m\tau} \leq \int_0^{t_n} e^{C_0 s}\,ds = C_0^{-1}\left(e^{C_0 t_n} - 1\right).$$

Combining the last two inequalities yields

$$\|u_h^n\|_V^2 + 2\tau\sum_{m=0}^{n-1}(u_h^m, A_h u_h^m)_V \leq e^{C_0 t_n}\left(\|u_h^0\|_V^2 + 3\tau\sum_{m=0}^{n-1}\|g_h^m\|_V^2\right),$$

whence (4.47) follows from the skew-adjointness of A_h^{cf} and (4.48) from the dissipative property of A_h^{upw}, see (3.64) and (3.65). □

RK2 and RK3 methods We proceed with the higher order methods. For RK2 schemes we can prove the stability analogously to RK1 schemes except that we can weaken the 2-CFL condition to the 4/3-CFL condition. The stability proof of RK3 schemes requires more attention but this is paid off with the necessity of only the usual CFL condition. For the sake of clarity we omit from now writing out the constants and use a generic constant C which is independet of τ and h but can depend on g, u, the constants in Lemmata 3.28, 3.32, 3.35, the RK coefficients and the CFL constants. The value of C can change at each occurance.

We begin by a short technical lemma.

Lemma 4.22 *(Bounds for U_h^{n1}, U_h^{n2}, R_2^n, R_3^n and the Peano kernels).*

i) For U_h^{n1} and U_h^{n2} the following bounds hold,

$$\|U_h^{n1}\|_V^2 \leq C(\|u_h^n\|_V^2 + \|g_h^n\|_V^2), \tag{4.51}$$
$$\|U_h^{n2}\|_V^2 \leq C(\|u_h^n\|_V^2 + \|g_h^n\|_V^2 + \|\tau\partial_t g_h^n\|_V^2). \tag{4.52}$$

ii) For $x \in [0,1]$ the Peano kernels are bounded by

$$\sup_{s\in[t_n,t_{n+1}]} \left|\kappa_{x,p}^{(q)}\left(\frac{s-t_n}{\tau}\right)\right| \leq \frac{1}{(p-q)!}x^{p-q}, \qquad \forall q \leq p. \tag{4.53}$$

iii) Let $g_h \in C^2(0,T;V_h)$ in the RK2 case and $g_h \in C^3(0,T;V_h)$ in the RK3 case. Then, the remainder terms can be estimated by

$$\|R_2^n\|_V^2 \leq C\tau^3 \int_{t_n}^{t_{n+1}} \|\partial_{tt} g_h(s)\|_V^2 \, ds, \tag{4.54}$$

and

$$\|R_3^n\|_V^2 \leq C\tau^5 \int_{t_n}^{t_{n+1}} \|\partial_{ttt} g_h(s)\|_V^2 \, ds + C\tau^5 \int_{t_n}^{t_{n+1}} \|A_h(\partial_{tt} g_h(s))\|_V^2 \, ds. \tag{4.55}$$

Remark 4.23 Note that the estimate (4.54) together with the boundedness of π_h imply

$$\|R_2^n\|_V^2 \leq C\tau^3 \int_{t_n}^{t_{n+1}} \|\partial_{tt} g(s)\|_V^2 \, ds.$$

In contrast, we need to apply the boundedness of A_h (4.2) and the usual CFL condition to the second term in the estimate (4.55) in order to get

$$\|R_3^n\|_V^2 \leq C\tau^5 \int_{t_n}^{t_{n+1}} \|\partial_{ttt} g(s)\|_V^2 \, ds + C\tau^3 \int_{t_n}^{t_{n+1}} \|\partial_{tt} g(s)\|_V^2 \, ds.$$

We observe the reduction by a factor τ^2 in the second term. This does not spoil a stability result but causes problems for the convergence, cf. Section 4.5. ◇

Proof: i) Using the triangle inequality and Young's inequality in (4.29) we infer

$$\|U_h^{n1}\|_V^2 \le C(\|u_h^n\|_V^2 + \|\tau A_h u_h^n\|_V^2 + \|\tau g_h^n\|_V^2).$$

Furthermore, we apply the boundedness of A_h (4.2) and the usual CFL condition to the second term and $\tau^2 \le 1$ to the third,

$$\|U_h^{n1}\|_V^2 \le C(\|u_h^n\|_V^2 + \|g_h^n\|_V^2).$$

This proves (4.51). The bound (4.52) is shown analogously.
ii) Let $x \in [0,1]$. It holds

$$\left|\kappa_{x,p}^{(q)}\left(\frac{s-t_n}{\tau}\right)\right| = \frac{1}{(p-q)!}\left(\frac{t_n+\tau x-s}{\tau}\right)_+^{p-q},$$

which is a monotonically decreasing function on $[t_n, t_{n+1}]$ and thus takes its maximum for $s = t_n$.
iii) We use the bound on the Peano kernels (4.53) to infer

$$\|R_2^n\|_V^2 \le C\tau^2\left(\int_{t_n}^{t_{n+1}} \|\partial_{tt} g_h(s)\|_V \, ds\right)^2 \le C\tau^3 \int_{t_n}^{t_{n+1}} \|\partial_{tt} g_h(s)\|_V^2 \, ds,$$

where the second estimate is obtained with the Cauchy-Schwarz inequality. The bound for R_3^n is proven analogously. □

Now, we state the stability results for RK2 and RK3 methods.

Lemma 4.24 *(Stability for RK2). Let $\{u_h^n\}_n$ be a RK2 discretization of (4.1) and let*

$$\mathfrak{g}_2^n = \|g_h^n\|_V^2 + \|\tau \partial_t g_h^n\|_V^2 + \|R_2^n\|_V^2.$$

Then, under the 4/3-CFL condition (4.45), the following results hold:

i) In the centered fluxes case,

$$\|u_h^n\|_V^2 \le C\left(\|u_h^0\|_V^2 + \tau \sum_{m=0}^{n-1} \mathfrak{g}_2^m\right). \tag{4.56}$$

ii) In the upwind fluxes case,

$$\|u_h^n\|_V^2 + \tau \sum_{m=0}^{n-1} \left[|u_h^m|_S^2 + |U_h^{m1}|_S^2\right] \le C\left(\|u_h^0\|_V^2 + \tau \sum_{m=0}^{n-1} \mathfrak{g}_2^m\right). \tag{4.57}$$

Proof: We apply the Cauchy-Schwarz inequality and Young's inequality to the RK2 energy identity (4.32), which yields

$$\begin{aligned}\|u_h^{n+1}\|_V^2 &+ \tau(u_h^n, A_h u_h^n)_V + \tau(U_h^{n1}, A_h U_h^{n1})_V \\ &\le \|u_h^n\|_V^2 + C\tau(\|u_h^n\|_V^2 + \|U_h^{n1}\|_V^2) + C\tau(\|g_h^n\|_V^2 + \|\tau\partial_t g_h^n\|_V^2 + \|R_2^n\|_V^2) \\ &\quad + \frac{1}{4}\|\tau^2 A_h^2 u_h^n - \tau^2 A_h g_h^n + \tau^2 \partial_t g_h^n + 2\tau R_2^n\|_V^2. \end{aligned} \tag{4.58}$$

For the last term we use the triangle inequality, Young's inequality and the boundedness of A_h (4.2) to infer

$$\begin{aligned}\frac{1}{4}\|\tau^2 A_h^2 u_h^n - \tau^2 A_h g_h^n + \tau^2 \partial_t g_h^n + 2\tau R_2^n\|_V^2 \le{}& C_h^4 c_\infty^4 \tau^4 h^{-4} \|u_h^n\|_V^2 + C_h^2 c_\infty^2 \tau^2 h^{-2} \|\tau g_h^n\|_V^2 \\ &+ \|\tau^2 \partial_t g_h^n\|_V^2 + \|2\tau R_2^n\|_V^2.\end{aligned}$$

We see that the first term requires the 4/3-CFL condition whereas for the second term we can use the usual CFL condition. Then, we have

$$\frac{1}{4}\|\tau^2 A_h^2 u_h^n - \tau^2 A_h g_h^n + \tau^2 \partial_t g_h^n + 2\tau R_2^n\|_V^2 \leq C\tau\|u_h^n\|_V^2 + C\|\tau g_h^n\|_V^2 + \|\tau^2 \partial_t g_h^n\|_V^2 + \|2\tau R_2^n\|_V^2.$$

Inserting this inequality together with $\tau^2 \leq \tau$ and (4.51) into (4.58) yields

$$\begin{aligned}\|u_h^{n+1}\|_V^2 + &\tau(u_h^n, A_h u_h^n)_V + \tau(U_h^{n1}, A_h U_h^{n1})_V \\ &\leq \|u_h^n\|_V^2 + C\tau\|u_h^n\|_V^2 + C\tau(\|g_h^n\|_V^2 + \|\tau\partial_t g_h^n\|_V^2 + \|R_2^n\|_V^2).\end{aligned}$$

The rest of the proof proceeds analogously to the proof of Lemma 4.20 for the forward Euler method, i. e. by summing from 0 to $n-1$ and applying the discrete Gronwall lemma. This gives

$$\begin{aligned}\|u_h^n\|_V^2 + &\tau \sum_{m=0}^{n-1} \left[(u_h^m, A_h u_h^m)_V + (U_h^{m1}, A_h U_h^{m1})_V\right] \\ &\leq e^{Cn\tau}\left(\|u_h^0\|_V^2 + C\tau \sum_{m=0}^{n-1}\left[\|g_h^m\|_V^2 + \|\tau\partial_t g_h^m\|_V^2 + \|R_2^n\|_V^2\right]\right).\end{aligned}$$

Then, (4.56) follows by the skew-adjointness of A_h^{cf} and (4.57) follows by the dissipative property of A_h^{upw}. □

Lemma 4.25 *(Stability for RK3). Let $\{u_h^n\}_n$ be a RK3 discretization of (4.1) and let*

$$\mathfrak{g}_3^n = \|g_h^n\|_V^2 + \|\tau\partial_t g_h^n\|_V^2 + \|\tau^2\partial_{tt} g_h^n\|_V^2 + \|R_3^n\|_V^2.$$

i) In the centered fluxes case assume that the usual CFL condition (4.44) is satisfied with

$$\varrho_{cf} \leq \sqrt{\frac{3}{2}}\left(C_h^{\mathrm{cf}}\right)^{-1}.$$

Then, there holds

$$\|u_h^n\|_V^2 \leq C\left(\|u_h^0\|_V^2 + \tau\sum_{m=0}^{n-1}\mathfrak{g}_3^m\right). \tag{4.59}$$

ii) In the upwind fluxes case assume that the usual CFL condition (4.44) is satisfied with

$$\varrho_{upw} \leq \min\left(\sqrt{\frac{3}{4}}(C_h^{\mathrm{upw}})^{-1}, \frac{5}{154}C_{\mathrm{bnd}}^{-2}\right).$$

Then, there holds

$$\|u_h^n\|_V^2 + \tau\sum_{m=0}^{n-1}\left[\frac{1}{12}|u_h^m|_S^2 + \frac{1}{3}|U_h^{m1}|_S^2 + \frac{1}{12}|U_h^{m2}|_S^2\right] \leq C\left(\|u_h^0\|_V^2 + \tau\sum_{m=0}^{n-1}\mathfrak{g}_3^m\right). \tag{4.60}$$

Remark 4.26 We observe that the upwind case requires a stronger assumption on the parameter ϱ to infer stability. Revisiting the RK3 energy identity (4.37) this manifests in the term $\frac{1}{3}((\tau A_h u_h^n - \tau g_h^n), \tau A_h(\tau A_h u_h^n - \tau g_h^n))_V$ which vanishes in the centered fluxes case but is a non-negative quantity which needs to be balanced in the upwind fluxes case . ◇

Proof: We apply the Cauchy-Schwarz inequality and Young's inequality to the RK3 energy identity (4.37), which yields

$$\begin{aligned}\|u_h^{n+1}\|_V^2+\tau(u_h^n,A_hu_h^n)_V+\frac{1}{3}\tau(U_h^{n1},A_hU_h^{n1})_V\\ +\frac{2}{3}\tau(U_h^{n2},A_hU_h^{n2})_V+\frac{1}{12}\|\tau^2A_h^2u_h^n-\tau^2A_hg_h^n+\tau^2\partial_tg_h^n\|_V^2\\ \le\|u_h^n\|_V^2+\frac{1}{3}\tau\big((\tau A_hu_h^n-\tau g_h^n),A_h(\tau A_hu_h^n-\tau g_h^n)\big)_V\\ +C\tau(\|u_h^n\|_V^2+\|U_h^{n1}\|_V^2+\|U_h^{n2}\|_V^2)+C\tau(\|g_h^n\|_V^2+\|\tau\partial_tg_h^n\|_V^2+\|\tau^2\partial_{tt}g_h^n\|_V^2+\|R_3^n\|_V^2)\\ +\frac{1}{36}\|\tau^3A_h^3u_h^n-\tau^3A_h^2g_h^n+\tau^3A_h\partial_tg_h^n-\tau^3\partial_{tt}g_h^n-6\tau R_3^n\|_V^2.\end{aligned}\tag{4.61}$$

We split the last term using the triangle inequality and Young's inequality into

$$\frac{1}{18}\|\tau^3A_h^3u_h^n-\tau^3A_h^2g_h^n+\tau^3A_h\partial_tg_h^n\|_V^2+\frac{1}{18}\|\tau^3\partial_{tt}g_h^n+6\tau R_3^n\|_V^2.\tag{4.62}$$

Furthermore, we use the boundedness property (4.2) of A_h and the usual CFL condition to infer

$$\frac{1}{18}\|\tau^3A_h^3u_h^n-\tau^3A_h^2g_h^n+\tau^3A_h\partial_tg_h^n\|\le\frac{1}{18}C_h^2\varrho^2\|\tau^2A_h^2u_h^n-\tau^2A_hg_h^n+\tau^2\partial_tg_h^n\|_V^2.\tag{4.63}$$

The triangle inequality and Young's inequality together with (4.63) show that (4.62) can be bounded by

$$\frac{1}{18}C_h^2\varrho^2\|\tau^2A_h^2u_h^n-\tau^2A_hg_h^n+\tau^2\partial_tg_h^n\|_V^2+C\tau(\|\tau^2\partial_{tt}g_h^n\|_V^2+\|R_3^n\|_V^2).\tag{4.64}$$

Inserting (4.64) into (4.61) and using the bounds (4.51) and (4.52) on $\|U_h^{n1}\|_V^2$ and $\|U_h^{n2}\|_V^2$ in (4.61) give

$$\begin{aligned}\|u_h^{n+1}\|_V^2+\tau(u_h^n,A_hu_h^n)_V+\frac{1}{3}\tau(U_h^{n1},A_hU_h^{n1})_V\\ +\frac{2}{3}\tau(U_h^{n2},A_hU_h^{n2})_V+\frac{1}{12}\|\tau^2A_h^2u_h^n-\tau^2A_hg_h^n+\tau^2\partial_tg_h^n\|_V^2\\ \le\|u_h^n\|_V^2+\frac{1}{3}\tau\big((\tau A_hu_h^n-\tau g_h^n),A_h(\tau A_hu_h^n-\tau g_h^n)\big)_V\\ +C\tau\|u_h^n\|_V^2+C\tau(\|g_h^n\|_V^2+\|\tau\partial_tg_h^n\|_V^2+\|\tau^2\partial_{tt}g_h^n\|_V^2+\|R_3^n\|_V^2)\\ +\frac{1}{18}C_h^2\varrho^2\|\tau^2A_h^2u_h^n-\tau^2A_hg_h^n+\tau^2\partial_tg_h^n\|_V^2.\end{aligned}\tag{4.65}$$

i) Centered fluxes case: Owing to the skew-ajointness of the centered fluxes operator A_h^{cf}, see (3.64), we can considerably simplify (4.65). In fact, there holds

$$\begin{aligned}\|u_h^{n+1}\|_V^2+\frac{1}{12}\|\tau^2(A_h^{\mathrm{cf}})^2u_h^n-\tau^2A_h^{\mathrm{cf}}g_h^n+\tau^2\partial_tg_h^n\|_V^2\\ \le\|u_h^n\|_V^2+C\tau\|u_h^n\|_V^2+C\tau(\|g_h^n\|_V^2+\|\tau\partial_tg_h^n\|_V^2+\|\tau^2\partial_{tt}g_h^n\|_V^2+\|R_3^n\|_V^2)\\ +\frac{1}{18}(C_h^{\mathrm{cf}})^2\varrho^2\|\tau^2(A_h^{\mathrm{cf}})^2u_h^n-\tau^2A_h^{\mathrm{cf}}g_h^n+\tau^2\partial_tg_h^n\|_V^2.\end{aligned}$$

Requiring $\varrho\le\sqrt{\frac{3}{2}}(C_h^{\mathrm{cf}})^{-1}$ enables us to balance the last term with its counterpart on the LHS. Then, we have

$$\|u_h^{n+1}\|_V^2-\|u_h^n\|_V^2\le C\tau\|u_h^n\|_V^2+C\tau(\|g_h^n\|_V^2+\|\tau\partial_tg_h^n\|_V^2+\|\tau^2\partial_{tt}g_h^n\|_V^2+\|R_3^n\|_V^2).$$

Summing this inequality from 0 to $n-1$ and subsequently applying the discrete Gronwall lemma implies assertion (4.59).

ii) Upwind fluxes case: We consider again equation (4.65) but this time we use the dissipative property (3.65) of the upwind fluxes operator A_h^{upw}. This yields

$$\begin{aligned}\|u_h^{n+1}\|_V^2+\tau|u_h^n|_S^2 + \frac{1}{3}\tau|U_h^{n1}|_S^2 + \frac{2}{3}\tau|U_h^{n2}|_S^2 + \frac{1}{12}\|\tau^2(A_h^{\mathrm{upw}})^2u_h^n - \tau^2A_h^{\mathrm{upw}}g_h^n + \tau^2\partial_t g_h^n\|_V^2\\ \leq \|u_h^n\|_V^2 + C\tau\|u_h^n\|_V^2 + \frac{1}{3}\tau|\tau A_h^{\mathrm{upw}}u_h^n - \tau g_h^n|_S^2\\ + C\tau(\|g_h^n\|_V^2 + \|\tau\partial_t g_h^n\|_V^2 + \|\tau^2\partial_{tt}g_h^n\|_V^2 + \|R_3^n\|_V^2)\\ + \frac{1}{18}(C_h^{\mathrm{upw}})^2\varrho^2\|\tau^2(A_h^{\mathrm{upw}})^2u_h^n - \tau^2A_h^{\mathrm{upw}}g_h^n + \tau^2\partial_t g_h^n\|_V^2. \end{aligned} \tag{4.66}$$

In contrast to the centered fluxes case we now have to additionally balance the term $\frac{1}{3}\tau|\tau A_h u_h^n - \tau g_h^n|_S^2$. We have $\tau A_h u_h^n - \tau g_h^n = u_h^n - U_h^{n1}$ and thus could estimate it by the weighted Young's inequality as

$$\frac{1}{3}\tau|\tau A_h u_h^n - \tau g_h^n|_S^2 \leq \frac{1}{3}\tau(1+\gamma)|u_h^n|_S^2 + \frac{1}{3}\tau(1+\gamma^{-1})|U_h^{n1}|_S^2,$$

for a positive number γ. But, by revisiting (4.66) we realize that we cannot cancel $\frac{1}{3}\tau(1+\gamma^{-1})|U_h^{n1}|_S^2$ for any choice of γ. Thus, we additionally incorporate the term U_h^{n2} by

$$\begin{aligned}u_h^n - U_h^{n1} &= (u_h^n - U_h^{n2}) + (U_h^{n2} - U_h^{n1})\\ &= (u_h^n - U_h^{n2}) + \frac{1}{2}\left(\tau^2(A_h^{\mathrm{upw}})^2u_h^n - \tau^2A_h^{\mathrm{upw}}g_h^n + \tau^2\partial_t g_h^n\right).\end{aligned}$$

Consequently, we get with the weighted Young's inequality

$$\begin{aligned}\frac{1}{3}\tau|\tau A_h u_h^n - \tau g_h^n|_S^2 \leq& \frac{1}{3}\tau(1+\gamma_1)|u_h^n - U_h^{n2}|_S^2\\ &+ \frac{1}{12}\tau(1+\gamma_1^{-1})\left|\tau^2(A_h^{\mathrm{upw}})^2u_h^n - \tau^2A_h^{\mathrm{upw}}g_h^n + \tau^2\partial_t g_h^n\right|_S^2,\end{aligned}$$

for a positive number γ_1. Applying the weighted Young's inequality once more on the first term with $\gamma_2 > 0$ and using the bound (3.24) on the S-seminorm in the second term eventually gives

$$\begin{aligned}\frac{1}{3}\tau|\tau A_h u_h^n - \tau g_h^n|_S^2 \leq& \frac{1}{3}\tau(1+\gamma_1)(1+\gamma_2)|u_h^n|_S^2 + \frac{1}{3}\tau(1+\gamma_1)(1+\gamma_2^{-1})|U_h^{n2}|_S^2\\ &+ \frac{1}{12}C_{\mathrm{bnd}}^2c_\infty h^{-1}\tau(1+\gamma_1^{-1})\|\tau^2(A_h^{\mathrm{upw}})^2u_h^n - \tau^2A_h^{\mathrm{upw}}g_h^n + \tau^2\partial_t g_h^n\|_V^2.\end{aligned} \tag{4.67}$$

In view of the LHS of (4.66) we choose $\gamma_1 = \frac{5}{72}$ and $\gamma_2 = \frac{11}{7}$ so that $\frac{1}{3}(1+\gamma_1)(1+\gamma_2) = \frac{11}{12}$ and $\frac{1}{3}(1+\gamma_1)(1+\gamma_2^{-1}) = \frac{7}{12}$. Furthermore, we apply the usual CFL condition to the last term of (4.67) and get

$$\begin{aligned}\frac{1}{3}\tau|\tau A_h u_h^n - \tau g_h^n|_S^2 \leq& \frac{11}{12}\tau|u_h^n|_S^2 + \frac{7}{12}|U_h^{n2}|_S^2 + \frac{77}{60}C_{\mathrm{bnd}}^2\varrho\|\tau^2(A_h^{\mathrm{upw}})^2u_h^n - \tau^2A_h^{\mathrm{upw}}g_h^n + \tau^2\partial_t g_h^n\|_V^2\end{aligned} \tag{4.68}$$

Finally we insert (4.68) into (4.66), which unfolds

$$\begin{aligned}\|u_h^{n+1}\|_V^2+\frac{1}{12}\tau|u_h^n|_S^2 + \frac{1}{3}\tau|U_h^{n1}|_S^2 + \frac{1}{12}\tau|U_h^{n2}|_S^2 + \frac{1}{12}\|\tau^2(A_h^{\mathrm{upw}})^2u_h^n - \tau^2A_h^{\mathrm{upw}}g_h^n + \tau^2\partial_t g_h^n\|_V^2\\ \leq \|u_h^n\|_V^2 + C\tau\|u_h^n\|_V^2 + C\tau\left(\|g_h^n\|_V^2 + \|\tau\partial_t g_h^n\|_V^2 + \|\tau^2\partial_{tt}g_h^n\|_V^2 + \|R_3^n\|_V^2\right)\\ + \left(\frac{1}{18}(C_h^{\mathrm{upw}})^2\varrho^2 + \frac{77}{60}C_{\mathrm{bnd}}^2\varrho\right)\|\tau^2(A_h^{\mathrm{upw}})^2u_h^n - \tau^2A_h^{\mathrm{upw}}g_h^n + \tau^2\partial_t g_h^n\|_V^2.\end{aligned}$$

By choosing $\varrho \leq \min\left(\sqrt{\frac{3}{4}}(C_h^{\mathrm{upw}})^{-1}, \frac{5}{154}C_{\mathrm{bnd}}^{-2}\right)$ we get rid of the last term and obtain

$$\begin{aligned}\|u_h^{n+1}\|_V^2+\frac{1}{12}\tau|u_h^n|_S^2 + \frac{1}{3}\tau|U_h^{n1}|_S^2 + \frac{1}{12}\tau|U_h^{n2}|_S^2\\ \leq \|u_h^n\|_V^2 + C\tau\|u_h^n\|_V^2 + C\tau(\|g_h^n\|_V^2 + \|\tau\partial_t g_h^n\|_V^2 + \|\tau^2\partial_{tt}g_h^n\|_V^2 + \|R_3^n\|_V^2).\end{aligned}$$

The assertion now follows by summing from 0 to $n-1$ and applying the discrete Gronwall lemma. □

4.5 Convergence

In this section we prove that an s-stage RK method applied to the semi-discrete evolution equation (4.1) yields a convergent approximation of Maxwell's equations (1.35). In Chapter 3 we have already proven the convergence of the semi-discretization (4.1) with order h^k and $h^{k+1/2}$ for centered and upwind fluxes, respectively. So far, we have shown the stability of the full-discretization and we recall that for the semi-discrete case stability already ensured convergence of suboptimal order h^k. We prove that this stays true for the fully discrete case. In fact, we show that we can directly deduce the convergence of order $h^k + \tau^s$ from the stability results. Last, we improve this bound to the order $h^{k+1/2} + \tau^s$ for the upwind case.

4.5.1 Error Analysis

We recall that in Section 3.6.1 we have introduced the spatial discretization errors $e^{\mathrm{cf}}(t) = u(t) - u_h^{\mathrm{cf}}(t)$ and $e^{\mathrm{upw}}(t) = u(t) - u_h^{\mathrm{upw}}(t)$, where u is the exact solution of Maxwell's equations (1.35) and u_h^{cf}, u_h^{upw} are the solutions of the semi-discretizations (3.66) and (3.67). In addition, we recall that we splitted the errors into a projection error and a dG error, $e(t) = e_\pi(t) - e_h(t)$ with $e \in \{e^{\mathrm{cf}}, e^{\mathrm{upw}}\}$ and $e_h \in \{e_h^{\mathrm{cf}}, e_h^{\mathrm{upw}}\}$. In the following definition we transfer this quantities to the fully discrete case. We refer to the general semi-discrete evolution problem (4.1) and only distinguish the cases $A_h = A_h^{\mathrm{cf}}$ and $A_h = A_h^{\mathrm{upw}}$ when necessary.

Definition 4.27 (Error types). Let $u(t) \in V_\star$ denote the exact solution of (1.35) and $\{u_h^n\}_n$ denote the RK approximation of (4.1). We define the *full discretization error*

$$e^n := u(t_n) - u_h^n.$$

Furthermore, we split the error into two parts

$$e^n = e_\pi^n - e_h^n,$$

where e_π^n is the *projection error* at time t_n,

$$e_\pi^n := u(t_n) - \pi_h u(t_n),$$

and e_h^n is given as

$$e_h^n := u_h^n - \pi_h u(t_n).$$

In Lemma 3.32 we have proven the bound $\|e_\pi^n\|_V \leq Ch^{k+1}|u(t_n)|_{H^{k+1}(\mathcal{T}_h)^6}$ for the projection error provided the exact solution satisfies $u \in H^{k+1}(\mathcal{T}_h)^6$. We define $B_\pi := |u(t_n)|_{H^{k+1}(\mathcal{T}_h)^6}$ for later purpose. Obviously, this error does not depend on the time discretization nor on the spatial discretization scheme. It only depends on the choice of the discrete space V_h. Both time and spatial discretization errors are containded in e_h^n. Thus, our aim is to bound e_h^n. We recall our approach for the spatial discretization error. We showed that the error $e_h(t)$ is governed by the same discrete evolution equation as $u_h(t)$ but with the defect $A_h e_\pi(t)$ instead of a source term. This enabled us to use the stability result to infer convergence.

The error recusions for the full-discretization rely on Taylor expansions of the projection of the exact solution $\pi_h u(t)$ and the consistency property $A_h u(t) = \pi_h A u(t)$ of the discrete operators, see Proposition 3.28. Indeed, we see by the consistency property that the projection of the exact solution satisfies following evolution equation, cf. Lemma 3.33,

$$\partial_t \pi_h u(t) = -A_h u(t) + g_h(t). \tag{4.69}$$

Clearly, this yields

$$\partial_{tt} \pi_h u(t) = -A_h(\partial_t u(t)) + \partial_t g_h(t),$$

or equivlanetly,

$$\partial_{tt}\pi_h u(t) = -A_h(\partial_t e_\pi(t) + \partial_t \pi_h u(t)) + \partial_t g_h(t).$$

Using (4.69) in the above equation we get

$$\partial_{tt}\pi_h u(t) = A_h^2 u(t) - A_h(\partial_t e_\pi(t)) - A_h g_h(t) + \partial_t g_h(t). \tag{4.70}$$

Differentiating (4.69) twice w. r. t. t gives

$$\partial_{ttt}\pi_h u(t) = -A_h(\partial_{tt} e_\pi(t) + \partial_{tt} u(t)) + \partial_{tt} g_h(t),$$

and applying (4.70) yields

$$\partial_{ttt}\pi_h u(t) = -A_h^3 u(t) + A_h^2(\partial_t e_\pi(t)) - A_h(\partial_{tt} e_\pi(t)) + A_h^2 g_h(t) - A_h(\partial_t g_h(t)) + \partial_{tt} g_h(t). \tag{4.71}$$

We will need this identities to derive error recursions for the RK methods. We begin with the forward Euler method.

Forward Euler method

Lemma 4.28 *(Error recursion for RK1). Let $u \in C^2(0,T;V) \cap C(0,T;V_\star)$ denote the exact solution of (1.35) and $\{u_h^n\}_n$ denote the forward Euler approximation of the semi-discrete problem (4.1). Then, the following error recursion holds*

$$e_h^{n+1} = e_h^n - \tau A_h e_h^n + \tau A_h e_\pi^n + \tau D_1^n, \tag{4.72}$$

where the defect D_1^n is given as

$$D_1^n = -\int_{t_n}^{t_{n+1}} \kappa_{1,1}\left(\frac{s-t_n}{\tau}\right) \pi_h(\partial_{tt} u(s))\, ds.$$

We see that the error satisfies the RK1 recursion (4.13) where the source term is substituted by $\tau A_h e_\pi^n + \tau D_1^n$.

Proof: We write the exact solution as first order Taylor expansion,

$$u(t_{n+1}) = u(t_n) + \tau \partial_t u(t_n) + \tau \int_{t_n}^{t_{n+1}} \kappa_{1,1}\left(\frac{s-t_n}{\tau}\right) \partial_{tt} u(s)\, ds.$$

Projecting onto V_h then gives

$$\pi_h u(t_{n+1}) = \pi_h u(t_n) + \tau \partial_t \pi_h u(t_n) - \tau D_1^n.$$

Using the consistency equation (4.69) we see that this is equivalent to

$$\pi_h u(t_{n+1}) = \pi_h u(t_n) - \tau A_h u(t) + \tau g_h^n - \tau D_1^n. \tag{4.73}$$

Subtracting (4.73) from the RK1 recursion (4.13) yields

$$e_h^{n+1} = e_h^n + \tau A_h e^n + \tau D_1^n, \tag{4.74}$$

whence the assertion follows by splitting the error in $A_h e^n = A_h(e_\pi^n - e_h^n)$. □

RK2 methods We introduce errors similar to Definition 4.27 associated with U_h^{n1}.

Definition 4.29 Let $u(t) \in V_\star$ denote the exact solution of (1.35). We define

$$U_\pi^{n1} := \pi_h u(t_n) + \tau \partial_t \pi_h u(t_n).$$

Furthermore, we define the error E^{n1} as

$$E^{n1} := u(t_n) + \tau \partial_t u(t_n) - U_h^{n1},$$

and split it into two parts

$$E^{n1} := E_\pi^{n1} - E_h^{n1}.$$

The first error E_π^{n1} is given as

$$E_\pi^{n1} := u(t_n) + \tau \partial_t u(t_n) - U_\pi^{n1},$$

and the second error E_h^{n1} is given as

$$E_h^{n1} := U_h^{n1} - U_\pi^{n1}.$$

Using (4.29) and (4.69) we see that there holds

$$E_h^{n1} = u_h^n - \tau A_h u_h^n + g_h^n - \pi_h u(t_n) - \tau \partial_t \pi_h u(t_n) = u_h^n - \pi_h u(t_n) + \tau A_h u(t_n) - \tau A_h u_h^n.$$

Thus, we have

$$E_h^{n1} = e_h^n + \tau A_h e^n = e_h^n - \tau A_h e_h^n + \tau A_h e_\pi^n. \tag{4.75}$$

We see that E_h^{n1} is the error of the forward Euler approximation without the defect D_1^n. Now, we state the RK2 error recursion.

Lemma 4.30 *(Error recursion for RK2). Let $u \in C^3(0,T;V) \cap C(0,T;V_\star)$ denote the exact solution of (1.35) and $\{u_h^n\}_n$ denote a RK2 approximation of the semi-discrete problem (4.1). Then, the following error recursion holds*

$$e_h^{n+1} = E_h^{n1} + \frac{1}{2}\tau A_h (e_h^n - E_h^{n1}) + \frac{1}{2}\tau^2 A_h (\partial_t e_\pi^n) + \tau D_2^n + \tau R_2^n, \tag{4.76}$$

where the defect D_2^n is given as

$$D_2^n = -\tau \int_{t_n}^{t_{n+1}} \kappa_{1,2}\left(\frac{s-t_n}{\tau}\right) \pi_h(\partial_{ttt} u(s))\, ds.$$

Proof: We apply second order Taylor expansion to the exact solution,

$$u(t_{n+1}) = u(t_n) + \tau \partial_t u(t_n) + \frac{1}{2}\tau^2 \partial_{tt} u(t_n) + \tau^2 \int_{t_n}^{t_{n+1}} \kappa_{1,2}\left(\frac{s-t_n}{\tau}\right) \partial_{ttt} u(s)\, ds.$$

Projecting onto V_h and subsequently applying (4.70) yields

$$\pi_h u(t_{n+1}) = U_\pi^{n1} + \frac{1}{2}\tau^2 A_h^2 u(t_n) - \frac{1}{2}\tau^2 A_h g_h^n + \frac{1}{2}\tau^2 \partial_t g_h^n - \frac{1}{2}\tau^2 A_h (\partial_t e_\pi^n) - \tau D_2^n. \tag{4.77}$$

Recall that we have shown in (4.30) that the RK2 approximations satisfies the following recursion

$$u_h^{n+1} = U_h^{n1} + \frac{1}{2}\tau^2 A_h^2 u_h^n - \frac{1}{2}\tau^2 A_h g_h^n + \frac{1}{2}\tau \partial_t g_h^n + \tau R_2^n. \tag{4.78}$$

Using Definition 4.29 we conclude that the difference between (4.78) and (4.77) is

$$e_h^{n+1} = E_h^{n1} - \frac{1}{2}\tau^2 A_h^2 e^n + \frac{1}{2}\tau^2 A_h (\partial_t e_\pi^n) + \tau D_2^n + \tau R_2^n.$$

From (4.75) we see $\tau A_h e^n = E_h^{n1} - e_h^n$, whence we infer the claim. □

RK3 methods Similar to RK2 methods we define errors associated with U_h^{n2}.

Definition 4.31 Let $u(t) \in V_\star$ denote the exact solution of (1.35). We define

$$U_\pi^{n2} := U_\pi^{n1} + \frac{1}{2}\tau^2 \partial_{tt} \pi_h u(t_n).$$

In addition, we introduce the error E^{n2} as

$$E^{n2} := u(t_n) + \tau \partial_t u(t_n) + \frac{1}{2}\tau^2 \partial_{tt} u(t_n) - U_h^{n2},$$

and split it into two parts

$$E^{n2} = E_\pi^{n2} - E_h^{n2}.$$

The first error E_π^{n2} is given as

$$E_\pi^{n2} := u(t_n) + \tau \partial_t u(t_n) + \frac{1}{2}\tau^2 \partial_{tt} u(t_n) - U_\pi^{n2},$$

and the second error E_h^{n2} is given as

$$E_\pi^{n2} := U_h^{n2} - U_\pi^{n2}.$$

Recalling the definition of U_h^{n2} given in (4.33) we see that there holds

$$E_h^{n2} = U_h^{n1} + \frac{1}{2}\tau A_h (u_h^n - U_h^{n1}) + \frac{1}{2}\tau^2 \partial_t g_h^n - U_\pi^{n1} - \frac{1}{2}\tau^2 \partial_{tt} \pi_h u(t_n).$$

Plugging the definition of U_h^{n1}, i. e. (4.29), and the identity (4.70) into the above equation we get

$$\begin{aligned} E_h^{n2} =& U_h^{n1} - U_\pi^{n1} + \frac{1}{2}\tau^2 A_h^2 u_h^n - \frac{1}{2}\tau^2 A_h u(t_n) + \frac{1}{2}\tau^2 A_h (\partial_t e_\pi^n) \\ =& E_h^{n1} - \frac{1}{2}\tau^2 A_h^2 e^n + \frac{1}{2}\tau^2 A_h (\partial_t e_\pi^n) \\ =& E_h^{n1} + \frac{1}{2}\tau A_h (e_h^n - E_h^{n1}) + \frac{1}{2}\tau^2 A_h (\partial_t e_\pi^n), \end{aligned} \tag{4.79}$$

where the third equality is obtained by (4.75). We observe that E_h^{n2} is the error of the RK2 approximation without remainder term R_2^n and without defect D_2^n, see (4.76).

Lemma 4.32 *(Error recursion for RK3). Let $u \in C^4(0,T;V) \cap C(0,T;V_\star)$ denote the exact solution of (1.35) and let $\{u_h^n\}_n$ denote a RK3 approximation of the semi-discrete problem (4.1). Then, the following error recursion holds*

$$e_h^{n+1} = E_h^{n2} + \frac{1}{3}\tau A_h (E_h^{n1} - E_h^{n2}) + \frac{1}{6}\tau^3 A_h (\partial_{tt} e_\pi^n) + \tau D_3^n + \tau R_3^n, \tag{4.80}$$

where the defect D_3^n is given as

$$D_3^n = -\tau^2 \int_{t_n}^{t_{n+1}} \kappa_{1,3}\left(\frac{s - t_n}{\tau}\right) \pi_h (\partial_t^4 u(s))\, ds.$$

Proof: The third order Taylor approximation of $u(t_{n+1})$ is given by

$$u(t_{n+1}) = u(t_n) + \tau\partial_t u(t_n) + \frac{1}{2}\tau^2\partial_{tt}u(t_n) + \frac{1}{6}\tau^3\partial_{ttt}u(t_n) + \tau^3\int_{t_n}^{t_{n+1}} \kappa_{1,3}\left(\frac{s-t_n}{\tau}\right)\partial_t^4 u(s)\,ds.$$

We project onto V_h and afterwards apply (4.71),

$$\begin{aligned}\pi_h u(t_{n+1}) =& U_\pi^{n2} - \frac{1}{6}\tau^3 A_h^3 u(t_n) + \frac{1}{6}\tau^3 A_h^2 g_h^n - \frac{1}{6}\tau^3 A_h(\partial_t g_h^n) + \frac{1}{6}\tau^3\partial_{tt}g_h^n \\ &+ \frac{1}{6}\tau^3 A_h^2(\partial_t e_\pi^n) - \frac{1}{6}\tau^3 A_h(\partial_{tt}e_\pi^n) - \tau D_3^n.\end{aligned}$$

We know from Lemma 4.17 that the RK3 recursion can be stated as

$$u_h^{n+1} = U_h^{n2} - \frac{1}{6}\tau^3 A_h^3 u_h^n + \frac{1}{6}\tau^3 A_h^2 g_h^n - \frac{1}{6}\tau^3 A_h(\partial_t g_h^n) + \frac{1}{6}\tau^3\partial_{tt}g_h^n + \tau R_3^n.$$

Subtracting the last two equations and using Definition 4.31 gives

$$e_h^{n+1} = E_h^{n2} + \frac{1}{6}\tau^3 A_h^3 e^n - \frac{1}{6}\tau^3 A_h^2(\partial_t e_\pi^n) - \frac{1}{6}\tau^3 A_h(\partial_{tt}e_\pi^n) + \tau D_3^n + \tau R_3^n.$$

The above equation can be rewritten as

$$e_h^{n+1} = E_h^{n2} + \frac{1}{3}\tau A_h\left(\frac{1}{2}\tau^2 A_h^2 e^n - \frac{1}{2}\tau^2 A_h(\partial_t e_\pi^n)\right) - \frac{1}{6}\tau^3 A_h(\partial_{tt}e_\pi^n) + \tau D_3^n + \tau R_3^n.$$

Hence, the claim is obtained by (4.79). □

We end this section by giving a bound for the defects.

Lemma 4.33 *(Bound on defects). Under the respective assumptions in Lemmata 4.28, 4.30 and 4.32 the defects D_1^n, D_2^n and D_3^n can be bounded by*

$$\|D_s^n\|_V^2 \le C\tau^{2s-1}\int_{t_n}^{t_{n+1}} \|\partial_t^{s+1}u(s)\|_V^2\,ds, \qquad s = 1,2,3. \tag{4.81}$$

Proof: Let $s \in \{1,2,3\}$ and note that we can write the defects as

$$D_s^n = \tau^{s-1}\int_{t_n}^{t_{n+1}} \kappa_{1,s}\left(\frac{s-t_n}{\tau}\right)\pi_h(\partial_t^{s+1}u(s))\,ds.$$

According to (4.53) the Peano kernels are bounded by

$$\sup_{s\in(t_n,t_{n+1})}\left|\kappa_{1,s}\left(\frac{s-t_n}{\tau}\right)\right| = \frac{1}{s!}.$$

Thus, we can deduce with (3.59) and the Cauchy-Schwarz inequality that it holds

$$\|D_s^n\|_V^2 \le C\tau^{2s-2}\left(\int_{t_n}^{t_{n+1}} \|\partial_t^{s+1}u(s)\|_V\,ds\right)^2 \le C\tau^{2s-1}\int_{t_n}^{t_{n+1}} \|\partial_t^{s+1}u(s)\|_V^2\,ds.$$

□

4.5.2 Centered Fluxes Case

In the previous section we have proven that the error e_h^{n+1} satisfies the same type of recurion as the RK approximation u_h^{n+1}. This enables us to apply the stability results obtained in Section 4.4 to the error and allows to prove convergence of order $h^k + \tau^s$. Therefore, we consider only the centered fluxes case here. Thus, by writing A_h we mean throughout this section A_h^{cf}.

We begin by stating the stability results for the error.

Lemma 4.34 *(Stability for error). Let $\{e_h^n\}_n$ be the error sequence of an s-stage RK approximation of the semi-discrete problem (4.1).*

i) For the forward Euler method let the 2-CFL condition (4.46) be satisfied. Then, there holds

$$\|e_h^n\|_V^2 \le C\tau \sum_{m=0}^{n-1} \|A_h e_\pi^m + D_1^m\|_V^2. \tag{4.82}$$

ii) For a RK2 method let the 4/3-CFL condition (4.45) be satisfied. Then, there holds

$$\|e_h^n\|_V^2 \le C\tau \sum_{m=0}^{n-1} \left[\|A_h e_\pi^m\|_V^2 + \|\tau A_h(\partial_t e_\pi^m)\|_V^2 + \|D_2^m + R_2^m\|_V^2\right] \tag{4.83}$$

iii) For a RK3 method let the usual CFL condition (4.44) with $\varrho \le \sqrt{\frac{3}{2}}(C_h^{\mathrm{cf}})^{-1}$ be satisfied. Then, there holds

$$\|e_h^n\|_V^2 \le C\tau \sum_{m=0}^{n-1} \left[\|A_h e_\pi^m\|_V^2 + \|\tau A_h(\partial_t e_\pi^m)\|_V^2 + \|\tau^2 A_h(\partial_{tt} e_\pi^m)\|_V^2 + \|D_3^m + R_3^m\|_V^2\right] \tag{4.84}$$

Proof: This immediately follows from the error recursions and the stability results, see Lemmata 4.28-4.32 and Lemmata 4.20, 4.24, 4.25, and the fact that $e_h^0 = 0$. □

Now, we prove convergence starting with the forward Euler method.

Forward Euler method

Theorem 4.35 *(Convergence for RK1). Let $u \in C^2(0,T;V) \cap C(0,T;V_{\star,k+1})$ denote the exact solution of (1.35) and $\{u_h^n\}_n$ denote the forward Euler approximation of the semi-discrete problem (4.1). Then, under the 2-CFL condition (4.46), there holds*

$$\|e^n\|_V^2 \le C(\tau^2 B_1 + h^{2k} B_1' + h^{2k+2} B_\pi),$$

with

$$B_1 = \int_0^{t_n} \|\partial_{tt} u(s)\|_V^2 \, ds, \qquad B_1' = \tau \sum_{m=0}^{n-1} |u(t_m)|_{H^{k+1}(\mathcal{T}_h)^6}^2.$$

Proof: We use the triangle inequality and Young's inequality in the stability estimate (4.82) ton infer

$$\|e_h^n\|_V^2 \le C\tau \sum_{m=0}^{n-1} \left[\|A_h e_\pi^m\|_V^2 + \|D_1^m\|_V^2\right] \le C\left(\tau \sum_{m=0}^{n-1} \|A_h e_\pi^m\|_V^2 + \tau^2 B_1\right).$$

Thereby, the second inequality is obtained with Lemma 4.33. Furthermore, the boundedness of A_h and the bounds for the projection errors, see Theorem 4.1 and Lemma 3.32, imply

$$\|A_h e_\pi^m\|_V^2 \le Ch^{-2}\|e_\pi^m\|_V^2 \le Ch^{2k}|u(t_m)|_{H^{k+1}(\mathcal{T}_h)^6}^2, \tag{4.85}$$

and we deduce

$$\|e_h^n\|_V^2 \leq C(\tau^2 B_1 + h^{2k} B_1'). \tag{4.86}$$

Last, we split the full error using Young's inequality,

$$\|e^n\|_V^2 \leq C(\|e_h^n\|_V^2 + \|e_\pi^n\|_V^2) \leq C\|e_h^n\|_V^2 + Ch^{2k+2} B_\pi,$$

where the second inequality is obtained from Lemma 3.32. Inserting (4.86) yields the claim. □

Convergence of RK2 methods is shown analogously.

RK2 methods

Theorem 4.36 *(Convergence for RK2). Let $u \in C^3(0,T;V) \cap C^1(0,T;H^k(\mathcal{T}_h)^6) \cap C(0,T;V_{\star,k+1})$ denote the exact solution of (1.35) with source term $g \in C^2(0,T;V)$. Furthermore, let $\{u_h^n\}_n$ denote a RK2 approximation of the semi-discrete problem (4.1). Then, under the 4/3-CFL condition (4.45), there holds*

$$\|e^n\|_V^2 \leq C(\tau^4 B_2 + h^{2k} B_2' + h^{2k+2} B_\pi),$$

where

$$B_2 = \int_0^{t_n} \|\partial_{tt} g(s)\|_V^2 + \|\partial_{ttt} u(s)\|_V^2 \, ds,$$

and

$$B_2' = \tau \sum_{m=0}^{n-1} \left[|u(t_m)|_{H^{k+1}(\mathcal{T}_h)^6}^2 + |\partial_t u(t_m)|_{H^k(\mathcal{T}_h)^6}^2 \right].$$

Proof: We bound the three terms on the RHS of the stability result (4.83). In (4.85) we already derived a bound for the first term and Lemmata 4.33 and 4.22 provide bounds for the terms D_2^n and R_2^n. This yields

$$\|e_h^n\|_V^2 \leq C\left(\tau^4 B_2 + \tau \sum_{m=0}^{n-1} \left[h^{2k} |u(t_m)|_{H^{k+1}(\mathcal{T}_h)^6} + \|\tau A_h(\partial_t e_\pi^m)\|_V^2\right]\right).$$

The remaining term is bounded using the boundedness property (4.2) of A_h together with the CFL condition and Lemma 3.32,

$$\|\tau A_h(\partial_t e_\pi^m)\|_V^2 \leq C\|\partial_t e_\pi^m\|_V^2 \leq Ch^k |\partial_t u(t_n)|_{H^k(\mathcal{T}_h)^6}^2. \tag{4.87}$$

We conclude

$$\|e_h^n\|_V^2 \leq C(\tau^4 B_2 + h^{2k} B_2').$$

The bound for the full error is obtained analogously to the previous theorem. □

RK3 methods We revisit the stability result for the error of RK3 methods given in Lemma 4.34 and observe with Lemma 4.17 that the remainder τR_3^n contains a term involving $A_h(\partial_{tt} g_h)$. In Lemma 4.22 we have shown the bound

$$\|R_3^n\|_V^2 \leq C\tau^5 \int_{t_n}^{t_{n+1}} \|\partial_{ttt} g_h(s)\|_V^2 \, ds + C\tau^5 \int_{t_n}^{t_{n+1}} \|A_h(\partial_{tt} g_h(s))\|_V^2 \, ds,$$

and pointed out that requiring solely $g \in C^3(0,T;V)$ admits only the bound

$$\|R_3^n\|_V^2 \leq C\tau^5 \int_{t_n}^{t_{n+1}} \|\partial_{ttt} g(s)\|_V^2 \, ds + C\tau^3 \int_{t_n}^{t_{n+1}} \|\partial_{tt} g(s)\|_V^2 \, ds.$$

This would lead to an order reduction from τ^3 to τ^2. It is possible to avoid this reduction but we have to require more regularity from the source term.

Proposition 4.37 *(Regularity and bound for the source term). Let* $g \in C^3(0,T;V) \cap C^2(0,T;V_\star)$. *Then, there holds*

$$\|R_3^n\|_V^2 \le C\tau^5 \int_{t_n}^{t_{n+1}} \|\partial_{ttt} g(s)\|_V^2 \, ds + C\tau^5 \int_{t_n}^{t_{n+1}} |\partial_{tt} g(s)|_{H^1(\mathcal{T}_h)^6}^2 \, ds. \tag{4.88}$$

Remark 4.38 We recall that $V_\star = \mathcal{D}(A) \cap H^1(\mathcal{T}_h)^6$ and that a function $v = [H, E]^T \in \mathcal{D}(A)$ has zero tangential component on the boundary, i. e. $n \times E|_{\partial\Omega} = 0$. Thus, the regularity assumption in Proposition 4.37 does not only concern the smoothness of the source term but also requires a (weak) boundary condition from. ◇

Proof: The bound on the first term is clear. For the second term we use a result proven in [16, Theorem 6.2], namely that there holds for all $v \in V_\star$

$$\|A_h \pi_h v - \pi_h A v\|_V \le C |v|_{H^1(\mathcal{T}_h)^6}, \tag{4.89}$$

with a constant independent of h. This enables the following estimate

$$\begin{aligned} \|A_h(\partial_{tt} g_h(s))\|_V =& \|A_h(\pi_h \partial_{tt} g(s))\|_V \\ \le& \|A_h(\pi_h \partial_{tt} g(s)) - \pi_h A(\partial_{tt} g(s))\|_V + \|\pi_h A(\partial_{tt} g(s))\|_V \\ \le& C |\partial_{tt} g(s)|_{H^1(\mathcal{T}_h)^6} + \|A(\partial_{tt} g(s))\|_V. \end{aligned}$$

For the second term we can use the bounds obtained in Proposition 3.28, namely

$$\|A(\partial_{tt} g(s))\|_V \le C(\|\nabla_h \times \partial_{tt} g(s)\|_V + h^{-1/2} |\partial_{tt} g(s)|_S).$$

From $\partial_{tt} g \in V_\star$ it follows $|\partial_{tt} g(s)|_S = 0$, see (3.45), and clearly we have $\|\nabla_h \times \partial_{tt} g(s)\|_V \le |\partial_{tt} g(s)|_{H^1(\mathcal{T}_h)^6}$, which concludes the proof. □

Now we can prove convergence.

Theorem 4.39 *(Convergence for RK3). Let* $u \in C^4(0,T;V) \cap C^2(0,T;H^{k-1}(\mathcal{T}_h)^6) \cap C^1(0,T;H^k(\mathcal{T}_h)^6) \cap C(0,T;V_{\star,k+1})$ *denote the exact solution of (1.35) with source term* $g \in C^3(0,T;V) \cap C^2(0,T;V_\star)$. *Furthermore, let* $\{u_h^n\}_n$ *denote a RK3 approximation of the semi-discrete problem (4.1). Assume that the step size satisies the usual CFL condition (4.44) with*

$$\varrho \le \sqrt{\frac{3}{2}} (C_h^{\mathrm{cf}})^{-1}.$$

Then, there holds

$$\|e^n\|_V^2 \le C\left(\tau^6 B_3 + h^{2k} B_3' + h^{2k+2} B_\pi\right),$$

where

$$B_3 = \int_0^{t_n} |\partial_{tt} g(s)|_{H^1(\mathcal{T}_h)^6}^2 + \|\partial_{ttt} g(s)\|_V^2 + \|\partial_t^4 u(s)\|_V^2 \, ds,$$

and

$$B_3' = \tau \sum_{m=0}^{n-1} \left[|u(t_m)|_{H^{k+1}(\mathcal{T}_h)^6}^2 + |\partial_t u(t_m)|_{H^k(\mathcal{T}_h)^6}^2 + |\partial_{tt} u(t_m)|_{H^{k-1}(\mathcal{T}_h)^6}^2 \right].$$

Proof: We bound the four terms in the stability result (4.84). Therefore, we apply (4.85) to the first term, (4.87) to the second term and Propostion 4.37 and Lemma 4.22 to the last. This yields

$$\|e_h^n\|_V^2 \leq C\left(\tau^6 B_3 + \tau \sum_{m=0}^{n-1}\left[|u(t_m)|^2_{H^{k+1}(\mathcal{T}_h)^6} + |\partial_t u(t_m)|^2_{H^k(\mathcal{T}_h)^6} + \|\tau^2 A_h(\partial_{tt} e_\pi^m)\|_V^2\right]\right).$$

For the remaining term we apply the boundedness property (4.2) of A_h together with the usual CFL condition and subsequently Lemma 3.32 to bound the projection error. This gives

$$\|\tau^2 A_h(\partial_{tt} e_\pi^m)\|_V^2 \leq C\tau^2 \|\partial_{tt} e_\pi^m\|_V^2 \leq C\tau^2 h^{2k-2} |\partial_{tt} u(t_m)|^2_{H^{k-1}(\mathcal{T}_h)^6} \leq Ch^{2k} |\partial_{tt} u(t_m)|^2_{H^{k-1}(\mathcal{T}_h)^6},$$

whence

$$\|e_h^n\|_V^2 \leq C(\tau^6 B_3 + h^{2k} B_3').$$

The result for the full error is obtained similar to Theorem 4.35. □

4.5.3 Upwind Fluxes Case

Now, we turn to the upwind fluxes case. So, let throughout this section denote $A_h = A_h^{\mathrm{upw}}$. Our aim is to prove convergence of order $h^{k+1/2} + \tau^s$ which is not possible using the stability results from Lemma 4.34. Instead, we begin by the following energy identities for the errors.

Lemma 4.40 *(Energy identities for errors). Let $\{e_h^n\}_n$ denote the error sequence of an s-stage RK approximation of the semi-discrete problem (4.1). Then, following energy identities hold:*

i) For the forward Euler method:

$$\|e_h^{n+1}\|_V^2 - \|e_h^n\|_V^2 + 2\tau|e_h^n|_S^2 = 2\tau(e_h^n, A_h E_\pi^{11} + D_1^n)_V + \|\tau A_h e_h^n - \tau A_h e_\pi^n - \tau D_1^n\|_V^2. \tag{4.90}$$

ii) For RK2 methods:

$$\begin{aligned}\|e_h^{n+1}\|_V^2 - \|e_h^n\|_V^2 + \tau|e_h^n|_S^2 + \tau|E_h^{n1}|_S^2 =& \tau(e_h^n, A_h E_\pi^{21})_V + \tau(E_h^{n1}, A_h E_\pi^{22} + 2(D_2^n + R_2^n))_V \\ &+ \frac{1}{4}\|\tau^2 A_h^2 e_h^n - \tau^2 A_h^2 e_\pi^n + \tau^2 A_h(\partial_t e_\pi^n) + 2\tau(D_2^n + R_2^n)\|_V^2.\end{aligned} \tag{4.91}$$

iii) For RK3 methods:

$$\begin{aligned}\|e_h^{n+1}\|_V^2 - \|e_h^n\|_V^2 + \tau|e_h^n|_S^2 + \frac{1}{3}\tau|E_h^{n1}|_S^2 + \frac{2}{3}\tau|E_h^{n2}|_S^2 + \frac{1}{12}\|\tau^2 A_h^2 e_h^n - \tau^2 A_h^2 e_\pi^n + \tau^2 A_h(\partial_t e_\pi^n)\|_V^2 \\ = \frac{1}{3}\tau|\tau A_h e_h^n - \tau A_h e_\pi^n|_S^2 \\ + \tau(e_h^n, A_h E_\pi^{31})_V + \frac{1}{3}\tau(E_h^{n1}, A_h E_\pi^{32})_V + \frac{2}{3}\tau(E_h^{n2}, A_h E_\pi^{33} + 3(D_3^n + R_3^n))_V \\ + \frac{1}{36}\|\tau^3 A_h^3 e_h^n - \tau^3 A_h^3 e_\pi^n + \tau^3 A_h^2(\partial_t e_\pi^n) - \tau^3 A_h(\partial_{tt} e_\pi^n) - 6\tau(D_3^n + R_3^n)\|_V^2.\end{aligned} \tag{4.92}$$

Thereby, the projection errors are given as

i) $E_\pi^{11} := e_\pi^n,$

ii) $E_\pi^{21} := e_\pi^n, \qquad E_\pi^{22} := e_\pi^n + \tau\partial_t e_\pi^n,$

iii) $E_\pi^{31} := e_\pi^n + \frac{1}{3}\tau\partial_t e_\pi^n, \qquad E_\pi^{32} := e_\pi^n, \qquad E_\pi^{33} := e_\pi^n + \tau\partial_t e_\pi^n + \frac{1}{2}\tau^2\partial_{tt} e_\pi^n.$

Proof: This follows by applying the RK energy identities obtained in Lemmata 4.13, 4.16 and 4.18 to their respective error recursion stated in Lemmata 4.28, 4.30 and 4.32. □

We observe that all three methods are featured by a certain structure of errors. Each type needs a different handling.

Lemma 4.41 *Let $q \in \{0,1,2\}$ and assume that the exact solution of (1.35) satisfies $u \in C^q(0,T;H^{k+1-q}(\mathcal{T}_h)^6)$. Then, under the usual CFL condition (4.44), we have for every $\gamma_q > 0$ the bound*

$$\tau\big(e_h^n, A_h(\tau^q \partial_t^q e_\pi^n)\big)_V \le \gamma_q \tau |e_h^n|_S^2 + C\tau h^{2k+1} |\partial_t^q u(t_n)|^2_{H^{k+1-q}(\mathcal{T}_h)^6}. \tag{4.93}$$

Remark 4.42 The proof of this lemma is based on Lemma 3.35 from Section 3.6. It was exactly this lemma which enabled us to prove the better convergence rate $h^{k+1/2}$ for the upwind case in the semi-discrete case. ◇

Proof: From Lemma 3.35 we get

$$\tau\big(e_h^n, A_h(\tau^q \partial_t^q e_\pi^n)\big)_V \le C\tau h^{-1/2} |e_h^n|_S \|\tau^q \partial_t^q e_\pi^n\|_V \le C\tau^{1/2} |e_h^n|_S \|\tau^q \partial_t^q e_\pi^n\|_V,$$

where the second estimate is gained with the usual CFL condition. We continue by applying the weighted Young's inequality with $\gamma_q > 0$ and Lemma 3.32,

$$\tau\big(e_h^n, A_h(\tau^q \partial_t^q e_\pi^n)\big)_V \le \gamma_q \tau |e_h^n|_S^2 + C\|\tau^q \partial_t^q e_\pi^n\|_V^2 \le \gamma_q \tau |e_h^n|_S^2 + C\tau^{2q} h^{2k+2-2q} |\partial_t^q u(t_n)|^2_{H^{k+1-q}(\mathcal{T}_h)^6}.$$

The assertion now follows by applying the usual CFL condition. □

Lemma 4.43 *Let $s \in \{1,2,3\}$, $q \in \{0,\dots,s-1\}$ and assume that the exact solution of (1.35) satisfies $u \in C^q(0,T;H^{k+1-q}(\mathcal{T}_h)^6)$. Then, under the usual CFL condition (4.44), it holds*

$$\|\tau^s A_h^{s-q}(\partial_t^q e_\pi^n)\|_V^2 \le C\tau h^{2k+1} |\partial_t^q u(t_n)|^2_{H^{k+1-q}(\mathcal{T}_h)^6}. \tag{4.94}$$

Proof: We use the boundedness of A_h (4.2) and the bounds on the projection errors provided by Lemma 3.32 to infer

$$\|\tau^s A_h^{s-q}(\partial_t^q e_\pi^n)\|_V^2 \le \tau^{2s} h^{-2s+2q} h^{2k+2-2q} |\partial_t^q u(t_n)|^2_{H^{k+1-q}(\mathcal{T}_h)^6} = \tau^{2s} h^{2k+2-2s} |\partial_t^q u(t_n)|^2_{H^{k+1-q}(\mathcal{T}_h)^6},$$

whence the assertion follows with the usual CFL condition. □

This two Lemmata allow us to prove convergence for the forward Euler method and the RK2 methods.

Forward Euler method

Theorem 4.44 *(Convergence for RK1). Let $u \in C^2(0,T;V) \cap C(0,T;V_{\star,k+1})$ denote the exact solution of (1.35) and $\{u_h^n\}_n$ denote the forward Euler approximation of the semi-discrete problem (4.1). Then, under the 2-CFL condition (4.46), there holds*

$$\|e^n\|_V^2 + \tau \sum_{m=0}^{n-1} |e^m|_S^2 \le C(\tau^2 B_1 + h^{2k+1} B_1' + h^{2k+2} B_\pi), \tag{4.95}$$

where B_1, B_1' are defined in Theorem 4.35.

Proof: We bound the two terms on the RHS of the forward Euler energy identity (4.90). For the first term we use Lemma 4.41 with $\gamma_0 = 1/2$ to get

$$\begin{aligned} 2\tau(e_h^n, A_h E_\pi^{11} + D_1^n)_V \leq & \tau|e_h^n|_S^2 + C\tau h^{2k+1}|u(t_n)|^2_{H^{k+1}(\mathcal{T}_h)^6} + 2\tau(e_h^n, D_1^n)_V \\ \leq & \tau|e_h^n|_S^2 + \tau\|e_h^n\|_V^2 + \tau\|D_1^n\|_V^2 + C\tau h^{2k+1}|u(t_n)|^2_{H^{k+1}(\mathcal{T}_h)^6}, \end{aligned} \tag{4.96}$$

where the second inequality is obtained by the Cauchy-Schwarz inequality and Young's inequality. The second term can be splitted with the triangle inequality and Young's inequality into

$$\|\tau A_h e_h^n - \tau A_h e_\pi^n - \tau D_1^n\|_V^2 \leq C(\|\tau A_h e_h^n\|_V^2 + \|\tau A_h e_\pi^n\|_V^2 + \|\tau D_1^n\|_V^2).$$

Now, we apply the boundedness property (4.2) of A_h together with the 2-CFL condition to the first term, Lemma 4.43 to the second term and use $\tau \leq \tau^2$ in the third term. Alltogether, this gives

$$\|\tau A_h e_h^n - \tau A_h e_\pi^n - \tau D_1^n\|_V^2 \leq C\tau\|e_h^n\|_V^2 + C\tau\|D_1^n\|_V^2 + C\tau h^{2k+1}|u(t_n)|^2_{H^{k+1}(\mathcal{T}_h)^6}. \tag{4.97}$$

Inserting (4.96) and (4.97) in the energy identity (4.90) yields

$$\|e_h^{n+1}\|_V^2 - \|e_h^n\|_V^2 + \tau|e_h^n|_S^2 \leq C\tau\|e_h^n\|_V^2 + C\tau\|D_1^n\|_V^2 + C\tau h^{2k+1}|u(t_n)|_{H^{k+1}(\mathcal{T}_h)^6}.$$

Summing this iequality from 0 to $n-1$ and using $e_h^0 = 0$ and the bound on the defect (4.81) give

$$\|e_h^n\|_V^2 + \tau\sum_{m=0}^{n-1}|e_h^m|_S^2 \leq C\tau\sum_{m=0}^{n-1}\|e_h^m\|_V^2 + C(\tau^2 B_1 + h^{2k+1}B_1').$$

Hence, we infer by using the discrete Gronwall lemma analogously as in the proof of Lemma 4.20 that there holds

$$\|e_h^n\|_V^2 + \tau\sum_{m=0}^{n-1}|e_h^m|_S^2 \leq C(\tau^2 B_1 + h^{2k+1}B_1'). \tag{4.98}$$

This concludes the estimate for the error part e_h^n. For the projection error e_π^n we use Lemmata 3.32 and 3.24 together with the usual CFL condition to infer

$$\begin{aligned} \|e_\pi^n\|_V^2 + \tau\sum_{m=0}^{n-1}|e_\pi^m|_S^2 \leq & Ch^{2k+2}|u(t_n)|^2_{H^{k+1}(\mathcal{T}_h)^6} + Ch^{2k+1}\tau\sum_{m=0}^{n-1}|u(t_m)|^2_{H^{k+1}(\mathcal{T}_h)^6} \\ = & C(h^{2k+2}B_\pi + h^{2k+1}B_1'). \end{aligned} \tag{4.99}$$

The estimate (4.95) now follows with Young's inequality and the bounds (4.98) and (4.99). □

RK2 methods

Theorem 4.45 *(Convergence for RK2). Let $u \in C^3(0,T;V) \cap C^1(0,T;H^k(\mathcal{T}_h)^6) \cap C(0,T;V_{\star,k+1})$ denote the exact solution of (1.35) with source term $g \in C^2(0,T;V)$. Further let $\{u_h^n\}_n$ denote a RK2 approximation of the semi-discrete problem (4.1). Then, under the 4/3-CFL condition (4.45), there holds*

$$\|e^n\|_V^2 + \tau\sum_{m=0}^{n-1}\left[|e^m|_S^2 + |E^{m1}|_S^2\right] \leq C(\tau^4 B_2 + h^{2k+1}B_2' + h^{2k+2}B_\pi), \tag{4.100}$$

where B_2 and B_2' are defined in Theorem 4.36.

Proof: We bound the three terms in the RK2 energy identity (4.91). The first two terms are bounded by Lemma 4.41 with $\gamma_0 = \gamma_1 = 1/2$,

$$\tau(e_h^n, A_h E_\pi^{21})_V + \tau(E_h^{n1}, A_h E_\pi^{22} + 2(D_2^n + R_2^n))_V \leq \frac{1}{2}\tau|e_h^n|_S^2 + \frac{1}{2}\tau|E_h^{n1}|_S^2 + 2\tau(E_h^{n1}, D_2^n + R_2^n)_V + C\tau h^{2k+1}\left(|u(t_n)|^2_{H^{k+1}(\mathcal{T}_h)^6} + |\partial_t u(t_n)|^2_{H^k(\mathcal{T}_h)^6}\right).$$

Applying further the Cauchy-Schwarz inequality and Young's inequality show

$$\tau(e_h^n, A_h E_\pi^{21})_V + \tau(E_h^{n1}, A_h E_\pi^{22} + 2(D_2^n + R_2^n))_V \leq \frac{1}{2}\tau|e_h^n|_S^2 + \frac{1}{2}\tau|E_h^{n1}|_S^2 + C\tau\|E_h^{n1}\|_V^2 + C\tau\left(\|D_2^n\|_V^2 + \|R_2^n\|_V^2\right) + C\tau h^{2k+1}\left(|u(t_n)|^2_{H^{k+1}(\mathcal{T}_h)^6} + |\partial_t u(t_n)|^2_{H^k(\mathcal{T}_h)^6}\right).$$

Furthermore, the third term in (4.91) is splitted using Young's inequality,

$$\frac{1}{4}\|\tau^2 A_h^2 e_h^n - \tau^2 A_h^2 e_\pi^n + \tau^2 A_h(\partial_t e_\pi^n) + 2\tau(D_2^n + R_2^n)\|_V^2 \leq C(\|\tau^2 A_h^2 e_h^n\|_V^2 + \|\tau^2 A_h^2 e_\pi^n\|_V^2 + \|\tau^2 A_h(\partial_t e_\pi^n)\|_V^2 + \|\tau D_2^n\|_V^2 + \|\tau R_2^n\|_V^2).$$

Now, the first term is bounded using (4.2) together with the 4/3-CFL condition, the second and third term are bounded by Lemma 4.43 and for the last two terms $\tau^2 \leq \tau$ is used. This yields the bound

$$\frac{1}{4}\|\tau^2 A_h^2 e_h^n - \tau^2 A_h^2 e_\pi^n + \tau^2 A_h(\partial_t e_\pi^n) + 2\tau(D_2^n + R_2^n)\|_V^2 \leq C\tau\|e_h^n\|_V^2 + C\tau\left(\|D_2^n\|_V^2 + \|R_2^n\|_V^2\right) + C\tau h^{2k+1}\left(|u(t_n)|^2_{H^{k+1}(\mathcal{T}_h)^6} + |\partial_t u(t_n)|^2_{H^k(\mathcal{T}_h)^6}\right).$$

In addition, using the triangle inequality, Young's inequality, the CFL condition and Lemma 4.43 we infer

$$\tau\|E_h^{n1}\|_V^2 \leq C\tau\|e_h^n\|_V^2 + C\tau h^{2k+1}|u(t_n)|^2_{H^{k+1}(\mathcal{T}_h)^6}.$$

Alltogether, we obtain

$$\|e_h^{n+1}\|_V^2 - \|e_h^n\|_V^2 + \frac{1}{2}\tau|e_h^n|_S^2 + \frac{1}{2}\tau|E_h^{n1}|_S^2 \leq C\tau\|e_h^n\|_V^2 + C\tau(\|D_2^n\|_V^2 + \|R_2^n\|_V^2) + C\tau h^{2k+1}\left(|u(t_n)|^2_{H^{k+1}(\mathcal{T}_h)^6} + |\partial_t u(t_n)|^2_{H^k(\mathcal{T}_h)^6}\right).$$

Summing from 0 to $n-1$ and applying the discrete Gronwall lemma yields

$$\|e_h^n\|_V^2 + \tau\sum_{m=0}^{n-1}\left[\frac{1}{2}|e_h^m|_S^2 + \frac{1}{2}|E_h^{m1}|_S^2\right] \leq C(\tau^4 B_2 + h^{2k+1}B_2'). \tag{4.101}$$

The projection error part is bounded as in (4.99). Indeed, there holds

$$\|e_\pi^n\|_V^2 + \tau\sum_{m=0}^{n-1}\left[\frac{1}{2}|e_\pi^m|_S^2 + \frac{1}{2}|E_\pi^{m1}|_S^2\right] \leq C(h^{2k+2}B_\pi + h^{2k+1}B_2'). \tag{4.102}$$

Combining (4.101) and (4.102) concludes the proof. □

RK3 methods We end this chapter by proving the convergence for RK3 methods.

Theorem 4.46 *(Convergence for RK3). Let* $u \in C^4(0,T;V) \cap C^2(0,T;H^{k-1}(\mathcal{T}_h)^6) \cap C^1(0,T;H^k(\mathcal{T}_h)^6) \cap C(0,T;V_{\star,k+1})$ *denote the exact solution of (1.35) with source term* $g \in$

$C^3(0,T;V)\cap C^2(0,T;V_\star)$. *Furthermore, let* $\{u_h^n\}_n$ *denote a RK3 approximation of the semi-discrete problem (4.1). Assume that the step size satisies the usual CFL condition (4.44) with*

$$\varrho \le \min\left(\sqrt{\frac{3}{4}}(C_h^{\mathrm{upw}})^{-1}, \frac{5}{154}C_{\mathrm{bnd}}^{-2}\right).$$

Then, there holds

$$\|e^n\|_V^2 + \tau\sum_{m=0}^{n-1}\left[\frac{1}{24}|e^m|_S^2 + \frac{1}{6}|E^{m1}|_S^2 + \frac{1}{24}|E^{m2}|_S^2\right] \le C(\tau^6 B_3 + h^{2k+1}B_3' + h^{2k+2}B_\pi), \tag{4.103}$$

where B_3 *and* B_3' *are given in Theorem 4.39.*

Proof: We bound the terms on the RHS of the RK3 energy identity (4.92). Therefore, we proceed in two steps.
i) The first step is motivated by the stability proof of Lemma 4.25. We write the first term in (4.92) as

$$\frac{1}{3}\tau|\tau A_h e_h^n - \tau A_h e_\pi^n|_S^2 = \frac{1}{3}\tau|(e_h^n - E_h^{n2}) + \frac{1}{2}(\tau^2 A_h^2 e_h^n - \tau^2 A_h^2 e_\pi^n + \tau^2 A_h(\partial_t e_\pi^n))|_S^2.$$

Applying the triangle inequality and weighted Young's inequalities with $\widetilde{\gamma}_1 = 5/72$ and $\widetilde{\gamma}_2 = 11/7$ we can draw the following estimate,

$$\begin{aligned}\frac{1}{3}\tau|\tau A_h e_h^n - \tau A_h e_\pi^n|_S^2 \le& \frac{1}{3}\tau(1+\widetilde{\gamma}_1)(1+\widetilde{\gamma}_2)|e_h^n|_S^2 + \frac{1}{3}\tau(1+\widetilde{\gamma}_1)(1+\widetilde{\gamma}_2^{-1})|E_h^{n2}|_S^2\\ &+ \frac{1}{12}\tau(1+\widetilde{\gamma}_1^{-1})|\tau^2 A_h^2 e_h^n - \tau^2 A_h^2 e_\pi^n + \tau^2 A_h(\partial_t e_\pi^n)|_S^2\\ =&\frac{11}{12}\tau|e_h^n|_S^2 + \frac{7}{12}\tau|E_h^{n2}|_S^2 + \frac{77}{60}\tau|\tau^2 A_h^2 e_h^n - \tau^2 A_h^2 e_\pi^n + \tau^2 A_h(\partial_t e_\pi^n)|_S^2.\end{aligned}$$

The bound for the S-seminorm (3.56) combined with the usual CFL condition with $\varrho \le \frac{5}{154}C_{\mathrm{bnd}}^{-2}$ show that the last term is bounded by

$$\frac{77}{60}\tau|\tau^2 A_h^2 e_h^n - \tau^2 A_h^2 e_\pi^n + \tau^2 A_h(\partial_t e_\pi^n)|_S^2 \le \frac{1}{24}\|\tau^2 A_h^2 e_h^n - \tau^2 A_h^2 e_\pi^n + \tau^2 A_h(\partial_t e_\pi^n)\|_V^2.$$

Hence, we deduce

$$\frac{1}{3}\tau|\tau A_h e_h^n - \tau A_h e_\pi^n|_S^2 \le \frac{11}{12}\tau|e_h^n|_S^2 + \frac{7}{12}\tau|E_h^{n2}|_S^2 + \frac{1}{24}\|\tau^2 A_h^2 e_h^n - \tau^2 A_h^2 e_\pi^n + \tau^2 A_h(\partial_t e_\pi^n)\|_V^2.$$

Next, we use Young's inequality to split the last term in the energy identity (4.92) into

$$\frac{1}{18}\|\tau^3 A_h^3 e_h^n - \tau^3 A_h^3 e_\pi^n + \tau^3 A_h^2(\partial_t e_\pi^n)\|_V^2 + \frac{1}{18}\|\tau^3 A_h(\partial_{tt} e_\pi^n) + 6\tau(D_3^n + R_3^n)\|_V^2.$$

For the first term we use the boundedness property (4.2) of A_h together with the usual CFL condition with $\varrho \le \sqrt{\frac{3}{4}}(C_h^{\mathrm{upw}})^{-1}$ to infer

$$\frac{1}{18}\|\tau^3 A_h^3 e_h^n - \tau^3 A_h^3 e_\pi^n + \tau^3 A_h^2(\partial_t e_\pi^n)\|_V^2 \le \frac{1}{24}\|\tau^2 A_h^2 e_h^n - \tau^2 A_h^2 e_\pi^n + \tau^2 A_h(\partial_t e_\pi^n)\|_V^2.$$

In summary, we obtain

$$\begin{aligned}\|e_h^{n+1}\|_V^2 - \|e_h^n\|_V^2 &+ \frac{1}{12}\tau|e_h^n|_S^2 + \frac{1}{3}\tau|E_h^{n1}|_S^2 + \frac{1}{12}\tau|E_h^{n2}|_S^2\\ &= \tau(e_h^n, A_h E_\pi^{31})_V + \frac{1}{3}\tau(E_h^{n1}, A_h E_\pi^{32})_V + \frac{2}{3}\tau\left(E_h^{n2}, A_h E_\pi^{33} + 3(D_3^n + R_3^n)\right)_V\\ &+ \frac{1}{18}\|\tau^3 A_h(\partial_{tt} e_\pi^n) + 6\tau(D_3^n + R_3^n)\|_V^2.\end{aligned} \tag{4.104}$$

ii) The second part is similar to the proof of Theorem 4.45. We apply Lemma 4.41 to the first three terms on the RHS of (4.104) with $\gamma_0 = 1/24$, $\gamma_1 = 1/2$ and $\gamma_2 = 1/16$. This yields

$$\begin{aligned}\tau(e_h^n, A_h E_\pi^{31})_V + \frac{1}{3}\tau(E_h^{n1}, A_h E_\pi^{32})_V + \frac{2}{3}\tau(E_h^{n2}, A_h E_\pi^{33} + 3(D_3^n + R_3^n))_V \\ \leq \frac{1}{24}\tau|e_h^n|_S^2 + \frac{1}{6}\tau|E_h^{n1}|_S^2 + \frac{1}{24}\tau|E_h^{n2}|_S^2 + 2\tau(E_h^{n2}, D_3^n + R_3^n)_V \\ + C\tau h^{2k+1}\left(|u(t_m)|^2_{H^{k+1}(\mathcal{T}_h)^6} + |\partial_t u(t_m)|^2_{H^k(\mathcal{T}_h)^6} + |\partial_{tt} u(t_m)|^2_{H^{k-1}(\mathcal{T}_h)^6}\right). \quad (4.105)\end{aligned}$$

Plugging (4.105) into (4.104) unfolds

$$\begin{aligned}\|e_h^{n+1}\|_V^2 - \|e_h^n\|_V^2 + \frac{1}{24}\tau|e_h^n|_S^2 + \frac{1}{6}\tau|E_h^{n1}|_S^2 + \frac{1}{24}\tau|E_h^{n2}|_S^2 \\ \leq 2\tau(E_h^{n2}, D_3^n + R_3^n)_V + \frac{1}{18}\|\tau^3 A_h(\partial_{tt} e_\pi^n) + 6\tau(D_3^n + R_3^n)\|_V^2 \\ + C\tau h^{2k+1}\left(|u(t_m)|^2_{H^{k+1}(\mathcal{T}_h)^6} + |\partial_t u(t_m)|^2_{H^k(\mathcal{T}_h)^6} + |\partial_{tt} u(t_m)|^2_{H^{k-1}(\mathcal{T}_h)^6}\right). \quad (4.106)\end{aligned}$$

Next, we split the first two terms on the RHS of (4.106). For the first term we use the Cauchy-Schwarz inequality and Young's inequality, whereas for second term we apply the triangle inequality and Young's inequality. This yields

$$\begin{aligned}2\tau(E_h^{n2}, D_3^n + R_3^n)_V + \frac{1}{18}\|\tau^3 A_h(\partial_{tt} e_\pi^n) + 6\tau(D_3^n + R_3^n)\|_V^2 \\ \leq C\tau\|E_h^{n2}\|_V^2 + C\tau(\|D_3^n\|_V^2 + \|R_3^n\|_V^2) + C\|\tau^3 A_h(\partial_{tt} e_\pi^n)\|_V^2 \\ \leq C\tau\|E_h^{n2}\|_V^2 + C\tau(\|D_3^n\|_V^2 + \|R_3^n\|_V^2) + Ch^{2k+1}|\partial_{tt} u(t_m)|^2_{H^{k-1}(\mathcal{T}_h)}, \quad (4.107)\end{aligned}$$

where we used Lemma 4.43 in the second inequality. We continue by using the boundedness property (4.2) of A_h together with the usual CFL condition and Lemma 4.43 to infer

$$\tau\|E_h^{n2}\|_V^2 \leq C\tau\|e_h^n\|_V^2 + C\tau h^{2k+1}\left(|u(t_m)|^2_{H^{k+1}(\mathcal{T}_h)^6} + |\partial_t u(t_m)|^2_{H^k(\mathcal{T}_h)^6}\right). \quad (4.108)$$

Finally, inserting (4.107) and (4.108) into (4.106) gives

$$\begin{aligned}\|e_h^{n+1}\|_V^2 - \|e_h^n\|_V^2 + \frac{1}{24}\tau|e_h^n|_S^2 + \frac{1}{6}\tau|E_h^{n1}|_S^2 + \frac{1}{24}\tau|E_h^{n2}|_S^2 \\ \leq C\tau\|e_h^n\|_V^2 + C\tau(\|D_3^n\|_V^2 + \|R_3^n\|_V^2) \\ + C\tau h^{2k+1}\left(|u(t_m)|^2_{H^{k+1}(\mathcal{T}_h)^6} + |\partial_t u(t_m)|^2_{H^k(\mathcal{T}_h)^6} + |\partial_{tt} u(t_m)|^2_{H^{k-1}(\mathcal{T}_h)^6}\right).\end{aligned}$$

Summing from 0 to $n-1$ gives

$$\|e_h^n\|_V^2 + \tau\sum_{m=0}^{n-1}\left[\frac{1}{24}|e_h^n|_S^2 + \frac{1}{6}|E_h^{n1}|_S^2 + \frac{1}{24}|E_h^{n2}|_S^2\right] \leq C\tau\sum_{m=0}^{n-1}\|e_h^m\|_V^2 + C(\tau^6 B_3 + h^{2k+1}B_3'),$$

whence applying the discrete Gronwall lemma yields

$$\|e_h^n\|_V^2 + \tau\sum_{m=0}^{n-1}\left[\frac{1}{24}|e_h^n|_S^2 + \frac{1}{6}|E_h^{n1}|_S^2 + \frac{1}{24}|E_h^{n2}|_S^2\right] \leq C(\tau^6 B_3 + h^{2k+1}B_3').$$

The projection errors are bounded similar to (4.99) and (4.102). Indeed, there holds

$$\|e_\pi^n\|_V^2 + \tau\sum_{m=0}^{n-1}\left[\frac{1}{24}|e_\pi^n|_S^2 + \frac{1}{6}|E_\pi^{n1}|_S^2 + \frac{1}{24}|E_\pi^{n2}|_S^2\right] \leq C(h^{2k+2}B_\pi + h^{2k+1}B_3').$$

Hence, the claim follows with Young's inequality and the two above estimates. □

Chapter 5

Implementation and Numerical Results

The last chapter is dedicated to numerical experiments which illustrate the theoretical results gained in this thesis, in particular in Chapter 3 and 4. We begin by shortly giving an insight into aspects of implementation. Then, we turn to numerical examples and confirm three main results of this thesis. First, we confirm the energy identities. In fact, our numerical results demonstrate the conservative behaviour of the centered fluxes discretization and the dissipative behaviour of the upwind fluxes discretization. In addition, our results illustrate the anti-dissipative behaviour of RK2 methods and the dissipative behaviour of RK3 methods. Then, we turn to verifying the convergence results. Therefore, we first check the convergence of the semi-discrete scheme and subsequently turn towards the convergence of the full discretization. Our examples are based on the matlab code of Hesthaven and Warburton, see [8, Chapter 6], in a complemented version of Tomislav Pažur, see [16]. Additional examples with implicit RK methods and TE Maxwell's equations on a deformed domain can also be found in [16, Chapter 7].

5.1 Implementation of dG Methods

We start with the semi-discrete problem obtained in Chapter 3 by discretiziting Maxwell's equations with dG methods: We search $u_h(t) \in C^1(0,T;V_h)$ such that

$$m_h(\partial_t u_h(t), \varphi_h) + a_h(u_h(t), \varphi_h) = (g(t), \varphi_h)_V \qquad \forall \varphi_h \in V_h. \tag{5.1}$$

The bilinear form m_h is given in (3.26) and accords with

$$m_h(\partial_t u_h(t), \varphi_h) = (\partial_t u_h(t), \varphi_h)_V. \tag{5.2}$$

Furthermore, we have $a_h \in \{a_h^{\mathrm{cf}}, a_h^{\mathrm{upw}}\}$ depending on the choice of the flux; thereby, a_h^{cf} is given in (3.33) and a_h^{upw} in (3.41).

The problem (5.1) is accessible for computitional solving since it is set in the finite dimensional space V_h. Owing to this fact, we can choose a basis of V_h consisting of finitely many vectors. Let us denote this basis with $\mathcal{V}_h = \{\varphi_1, \dots, \varphi_N\}$. The dimension N of the space V_h is given in Section 2.1.3 and depends on the number of mesh elements and on the polynomial degree we work with in the dG discretization. Since the dG approximation $u_h(t)$ is an element of the space V_h we deduce that there is a (unique) *coefficient vector* $\mathbf{u_h}(t) = [u_{h,1}(t), \dots, u_{h,N}(t)]^T \in \mathbb{R}^N$ such that

$$u_h(t) = \sum_{m=1}^{N} u_{h,m}(t)\varphi_m. \tag{5.3}$$

Obviously, for equation (5.1) it is equivalent to hold for all $\varphi_h \in V_h$ or to hold for all basis functions $\varphi_m \in \mathcal{V}_h$. Following this idea and further employing (5.2) and (5.3) we deduce that the problem

$$\sum_{m=1}^{N} (\varphi_m, \varphi_l)_V \, u'_{h,m}(t) + \sum_{m=1}^{N} a_h(\varphi_m, \varphi_l) \, u_{h,m}(t) = (g(t), \varphi_l)_V \qquad \forall l = 1, \dots, N, \tag{5.4}$$

is equivalent to (5.1). This motivates the following definition.

Definition 5.1 (Mass and stiffness matrix). We define the *mass matrix* $\mathbf{M} \in \mathbb{R}^{N\times N}$ by

$$\mathbf{M} := \left[(\varphi_m, \varphi_l)_V\right]_{l,m=1}^{N}, \tag{5.5}$$

and the *stiffness matrix* $\mathbf{A} \in \mathbb{R}^{N\times N}$ by

$$\mathbf{A} := \left[a_h\left(\varphi_m, \varphi_l\right)\right]_{l,m=1}^{N}. \tag{5.6}$$

Owing to the choice of the space V_h, both matrices are sparse. Furthermore, the mass matrix is block diagonal and symmetric positive definite and thus invertible. For a deeper insight we refer to [17, Appendix A]. We continue by defining

$$\mathbf{g_h}(t) := \left[(g(t), \varphi_m)_V\right]_{m=1}^{N} \in \mathbb{R}^N,$$

and alltogether have derived the following equivalent formulation of (5.4)

$$\mathbf{M}\mathbf{u_h'}(t) + \mathbf{A}\mathbf{u_h}(t) = \mathbf{g_h}(t).$$

Let us shortly reveal the connection between the source term $g_h(t)$ in our discretizations and $\mathbf{g_h}(t)$. Recall that we have defined $g_h(t) = \pi_h g(t)$ and thus there holds

$$g_h(t) = \pi_h g(t) = \sum_{m=1}^{N} (g(t), \varphi_m)_V \varphi_m = \sum_{m=1}^{N} g_{h,m}(t)\varphi_m,$$

where $g_{h,m}(t)$ denotes the m-th component of $\mathbf{g_h}(t)$. As already commented, $\mathbf{M}$ is invertible and we conclude

$$\mathbf{u_h'}(t) + \mathbf{M}^{-1}\mathbf{A}\mathbf{u_h}(t) = \widetilde{\mathbf{g_h}}(t), \tag{5.7}$$

with $\widetilde{\mathbf{g_h}}(t) := \mathbf{M}^{-1}\mathbf{g_h}(t)$. It is crucial to compare this problem to (3.66), (3.67), or alternatively to (4.1), in order to realize that the matrix $\mathbf{M}^{-1}\mathbf{A}$ corresponds to the discrete operator A_h.

5.2 Numerical Results

We consider TM polarized Maxwell's equations (1.9) for our numerical experiments. We recall that in this special case Maxwell's equations read as

$$\partial_t \begin{bmatrix} H_x \\ H_y \\ E_z \end{bmatrix} + A_{TM} \begin{bmatrix} H_x \\ H_y \\ E_z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -J_z \end{bmatrix} \qquad \text{in } (0,T)\times\Omega, \tag{5.8}$$

where the TM-Maxwell operator is given as

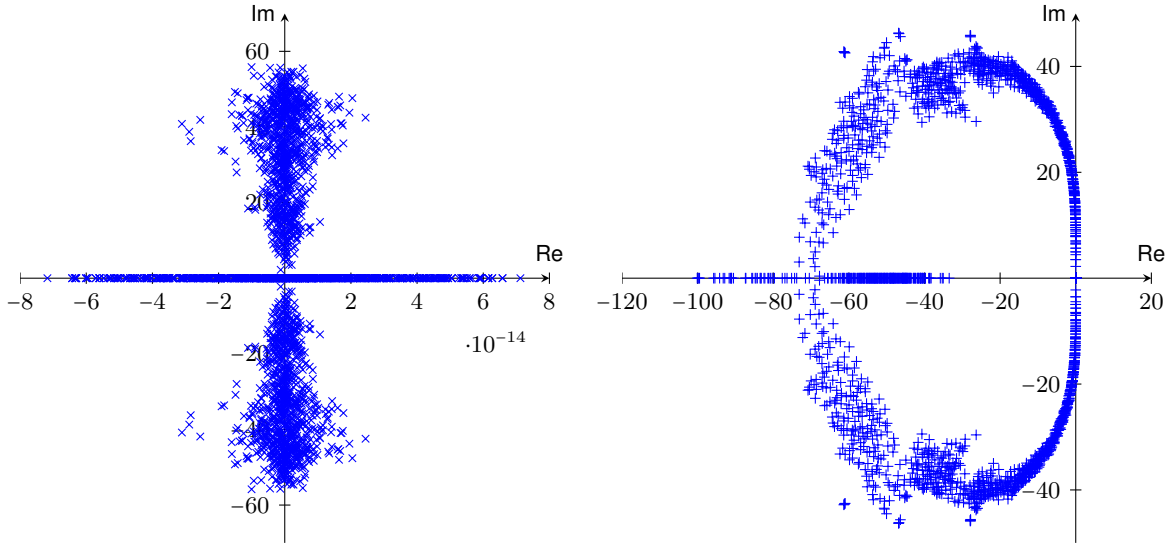
$$A_{TM} := \begin{bmatrix} 0 & 0 & \partial_y \\ 0 & 0 & -\partial_x \\ \partial_y & -\partial_x & 0 \end{bmatrix}. \tag{5.9}$$

We choose the domain as $\Omega = [-1,1]^2$ and the medium to be homogeneous, i. e. $\varepsilon = \mu = 1$.

Discretizing (5.8) in space yields a semi-discrete system as (5.7), i. e.

$$\begin{bmatrix} H_{h,x}' \\ H_{h,y}' \\ E_{h,z}' \end{bmatrix} + \mathbf{A_{h,TM}} \begin{bmatrix} H_{h,x} \\ H_{h,y} \\ E_{h,z} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\pi_h J_z \end{bmatrix}, \tag{5.10}$$

where $\mathbf{A_{h,TM}}$ corresponds to the discretization of the operator A_{TM}. The eigenvalues of $\mathbf{A_{h,TM}}$ for $k = 3$ and $h = 0.25$ are plotted in Figures 5.1(a) and 5.1(b). The first figure shows the eigenvalues when cenetred fluxes are used whereas for the second figures upwind fluxes were considered. We see that for upwind fluxes the eigenvalues lie in the left half-plan with a concentration towards to imaginary axis. In contrast, the eigenvalues for centered fluxes are close to the imaginary axis. Despite the occuring real parts are very small the eigenvalues with a positive real part can cause numerical instabilities.



(a) $h = 0.25$, $k = 3$, centered fluxes

(b) $h = 0.25$, $k = 3$, upwind fluxes

Figure 5.1: Eigenvalues of $\mathbf{A}_{\mathbf{h},\mathbf{TM}}$

5.2.1 Energy

We begin our numerical experiments with verifying the theorical results concering the evolution of the energy of a full discretization of (5.10) with explicit RK methods. We therefore consider an example of TM Maxwell's equations (5.8) given in [8, Chapter 6] without external forcing, $J_z = 0$, and with initial value

$$\begin{bmatrix} H_x(0) \\ H_y(0) \\ E_z(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \sin(\pi x)\sin(\pi y) \end{bmatrix}. \tag{5.11}$$

The exact solution is given by

$$\begin{bmatrix} H_x(t,x,y) \\ H_y(t,x,y) \\ E_z(t,x,y) \end{bmatrix} = \begin{bmatrix} -\frac{\pi}{\omega}\sin(\pi x)\cos(\pi y)\sin(\omega t) \\ \frac{\pi}{\omega}\cos(\pi x)\sin(\pi y)\sin(\omega t) \\ \sin(\pi x)\sin(\pi y)\cos(\omega t) \end{bmatrix}, \tag{5.12}$$

where $\omega := \sqrt{2}\pi$ is the resonance frequency. We can easily compute the norm of the solution as

$$\left\| \begin{bmatrix} H_x(t,\cdot,\cdot) \\ H_y(t,\cdot,\cdot) \\ E_z(t,\cdot,\cdot) \end{bmatrix} \right\|_V = 1, \qquad \forall t \geq 0.$$

In Figure 5.2 the energy of the full discrete solution at time $T = 10$ is plotted. We used either centered or upwind fluxes combined with the two- and three-stage Heun method as representant for RK2 and RK3 methods, respectively. For Figure 5.2(a) we used $h = 1$ and $k = 3$ and for Figure 5.2(b) we used $h = 0.5$ and $k = 4$ to illustrate the effect of a finer mesh and a higher polynomial degree on the energy. Owing to Theorem 3.29 and Remark 3.30 we expect that for upwind fluxes the latter discretization yields less dissipation which is clearly verified by our example. Furthermore, by the same theorem, we expect that the centered fluxes discretization is conservative. This is also confirmed if we consider Figures 5.2(a), 5.2(b) for small step sizes τ. This ensures that we deal with a negligible time discretization error and thus that the spatial discretization properties are dominant. Then, we see that the centered fluxes approximation yields energy equal to 1 which is in accordance with the exact solution.

In addition, Figure 5.2 approves the superior stability properties of the RK3 methods compared to RK2 methods which were reflected in the necessity of the stronger 4/3-CFL condition for the RK2 methods in contrast to the usual CFL condition for RK3 methods. Indeed, we see that RK3 methods allow a larger τ than RK2 methods. Furthermore, we observe that RK2 methods benefit from stabilization, i. e. upwind fluxes admit a larger limit for τ than the centered fluxes. This is explained by recalling the RK2 energy identity (4.22). For centered fluxes it reads as

$$\|u_h^{n+1}\|_V^2 = \|u_h^n\|_V^2 + \frac{1}{4}\|\tau^2 (A_h^{\mathrm{cf}})^2 u_h^n\|_V^2,$$

whereas for upwind fluxes we have

$$\|u_h^{n+1}\|_V^2 + \tau|u_h^n|_S^2 + \tau|U_h^{n1}|_S^2 = \|u_h^n\|_V^2 + \frac{1}{4}\|\tau^2 (A_h^{\mathrm{upw}})^2 u_h^n\|_V^2.$$

We see that the appearance of the two S-seminorm terms in the upwind case allow to compensate to some extent the anti-dissipative term $\frac{1}{4}\|\tau^2 (A_h^{\mathrm{upw}})^2 u_h^n\|_V^2$ which is not possible with centered fluxes. In contrary to RK2 methods working with stabilization requires a smaller limit for τ in the RK3 case. This is seen with the RK3 energy identity (4.25) which yields in the centered fluxes case

$$\|u_h^{n+1}\|_V^2 + \frac{1}{12}\|\tau^2 (A_h^{\mathrm{cf}})^2 u_h^n\|_V^2 = \|u_h^n\|_V^2 + \frac{1}{36}\|\tau^3 (A_h^{\mathrm{cf}})^3 u_h^n\|_V^2,$$

and in the upwind case

$$\begin{aligned}\|u_h^{n+1}\|_V^2 + \tau|u_h^n|_S^2 + \frac{1}{3}\tau|U_h^{n1}|_S^2 + \frac{2}{3}\tau|U_h^{n2}|_S^2 + \frac{1}{12}\|\tau^2 (A_h^{\mathrm{upw}})^2 u_h^n\|_V^2 \\ = \|u_h^n\|_V^2 + \frac{1}{3}\tau|\tau A_h^{\mathrm{upw}} u_h^n|_S^2 + \frac{1}{36}\|\tau^3 (A_h^{\mathrm{upw}})^3 u_h^n\|_V^2.\end{aligned}$$

We have proven in Lemma 4.25 that we can estimate the anti-dissipative term $\frac{1}{36}\|\tau^3 A_h^3 u_h^n\|_V^2$ by the term $\frac{1}{18} C_h^2 \varrho^2 \|\tau^2 A_h^2 u_h^n\|_V^2$. This enabled us to choose in the centered fluxes case $\varrho_{cf} \le \sqrt{\frac{3}{2}}(C_h^{\mathrm{cf}})^{-1}$ which even yields a dissipative behaviour

$$\|u_h^{n+1}\|_V \le \|u_h^n\|_V.$$

In the upwind fluxes case we could not balance the additional anti-dissipative term $\frac{1}{3}\tau|\tau A_h^{\mathrm{upw}} u_h^n|_S^2$ solely with the S-seminorm terms on the LHS but had to balance it with $\frac{1}{12}\|\tau^2 (A_h^{\mathrm{upw}})^2 u_h^n\|_V^2$, too. This led to a larger CFL coeffiecient $\varrho_{upw} \le \min\left(\sqrt{\frac{3}{4}}(C_h^{\mathrm{upw}})^{-1}, \frac{5}{154} C_{\mathrm{bnd}}^{-2}\right)$ which explains the smaller limit for τ compared to the centered fluxes case. However, if this limit is satisfied, RK3 methods in combination with upwind fluxes are also dissipative, see Figure 5.2.

5.2.2 Convergence of the Semi-Discretization

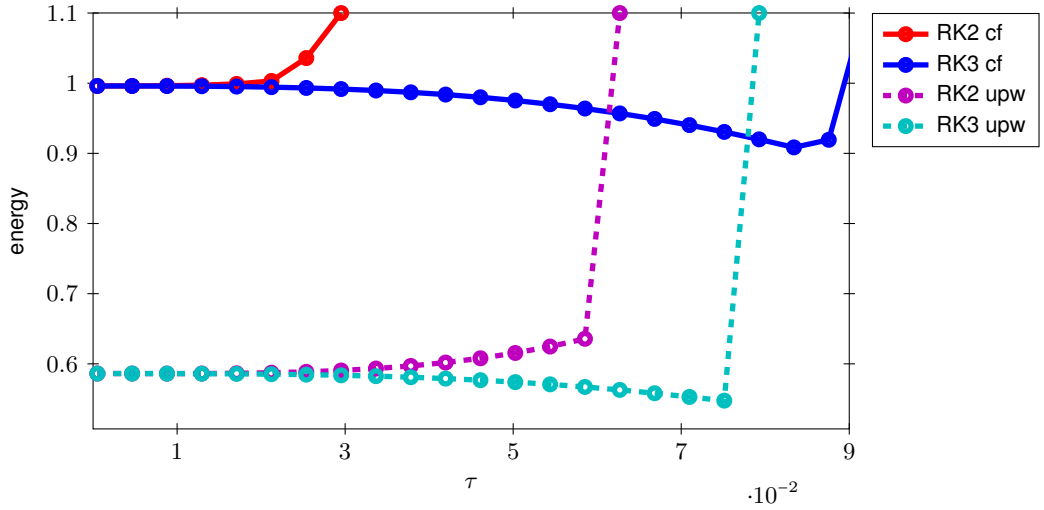
Now, let us check the convergence of the dG semi-discretization proven in Theorems 3.34 and 3.36. We adapt an example from [4] and use the following source term

$$J_z = -e^t\left[(1-x^2)(1-y^2) + 2(1-x^2) + 2(1-y^2)\right].$$

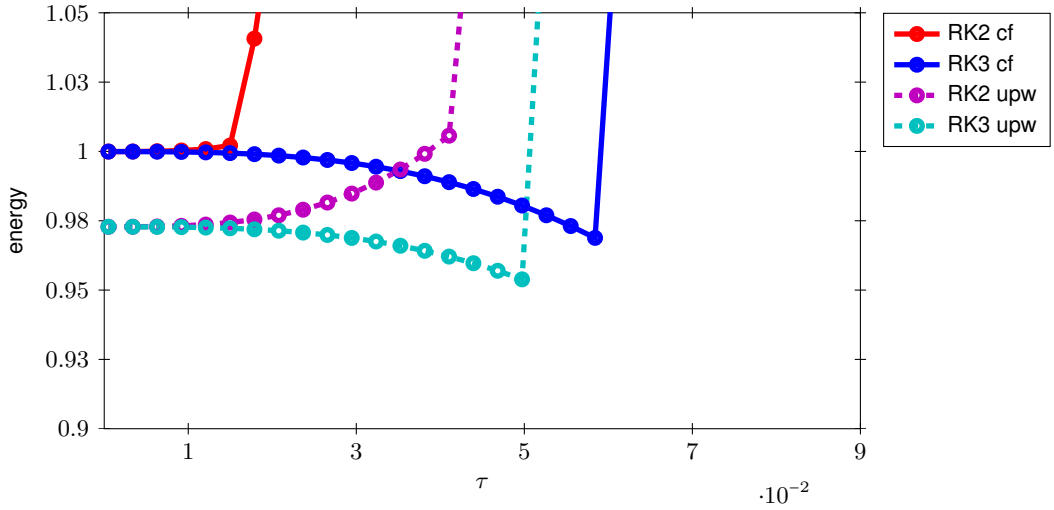
Then, the exact solution of (5.8) is given by

$$\begin{bmatrix} H_x(t,x,y) \\ H_y(t,x,y) \\ E_z(t,x,y) \end{bmatrix} = \begin{bmatrix} 2e^t(1-x^2)y \\ -2e^t x(1-y^2) \\ e^t(1-x^2)(1-y^2) \end{bmatrix}. \tag{5.13}$$

Since we want to investigate the spatial convergence we have to ensure that the time discretization error is small. Therefore, we choose a 3-stage Gauss collocation method for the time integration. This method is of order 6 and choosing $\tau = 0.01$ guarantees that the time discretization error is negligible. In Figure 5.3 the full error, i. e. $e^n = u(t_n) - u_h^n$, of both the centered fluxes and the upwind fluxes discretization is plotted. We observe convergence of order h^k in the centered fluxes case and h^{k+1} in the upwind fluxes case.

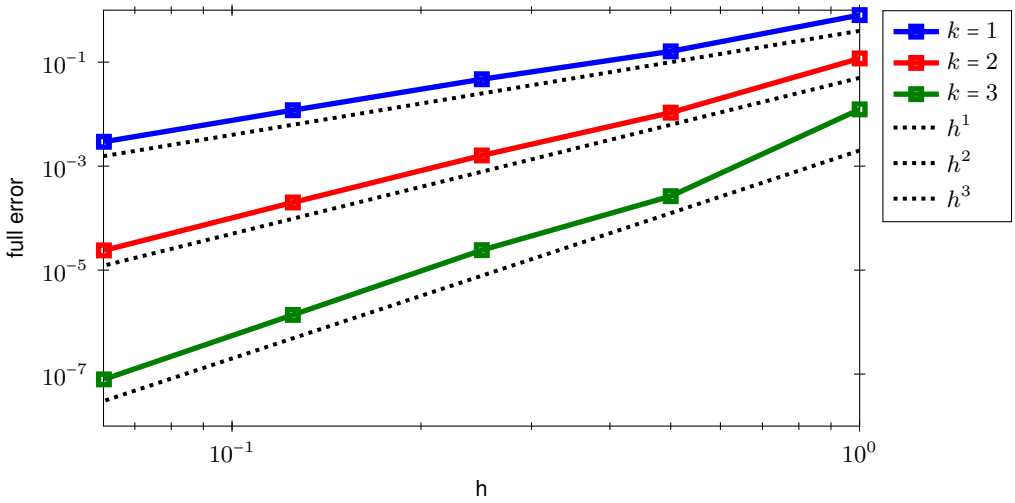


(a) meshsize $h = 1$, polynomial order $k = 3$

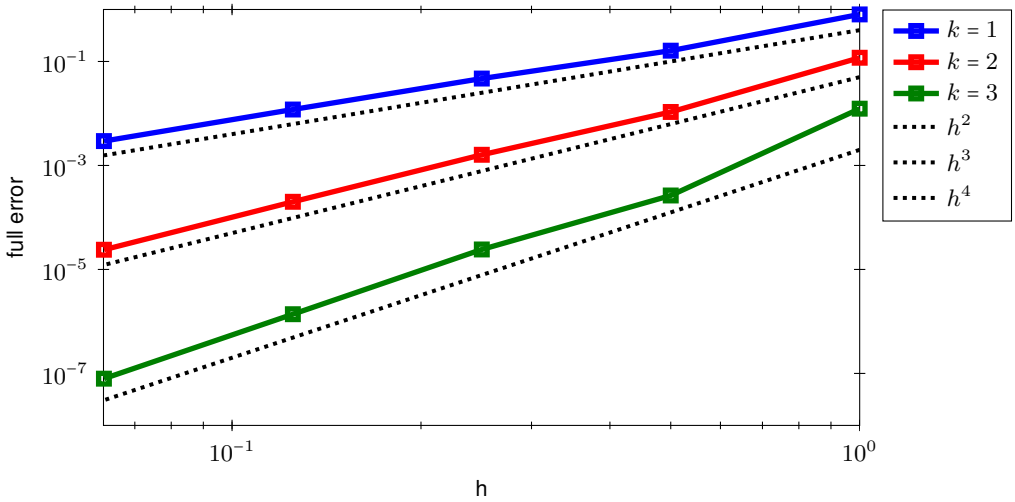


(b) meshsize $h = 0.5$, polynomial order $k = 4$

Figure 5.2: Electromagnetic energy at time $T = 10$



(a) centered fluxes



(b) upwind fluxes

Figure 5.3: Convergence of semi-discretization

5.2.3 Convergence of the Full Discretization

Finally, we illustrate the fully discrete convergence results given in Theorems 4.35, 4.36, 4.39 for RK1, RK2 and RK3 methods combined with centered fluxes and in Theorems 4.44, 4.45, 4.46 for RK1, RK2 and RK3 methods combined with upwind fluxes. In the previous section we already verified the convergence rate for the spatial discretization. Thus, we now check the time discretization results. We give a homogeneous and an inhomogeneous example.

Example 1 We begin by considering (5.8) without source term and with the initial value (5.11). Observe that the semi-discretization of (5.8) is

$$\begin{bmatrix} H'_{h,x}(t) \\ H'_{h,y}(t) \\ E'_{h,z}(t) \end{bmatrix} + \mathbf{A_{h,TM}} \begin{bmatrix} H_{h,x}(t) \\ H_{h,y}(t) \\ E_{h,z}(t) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

and consequently the exact solution is given by

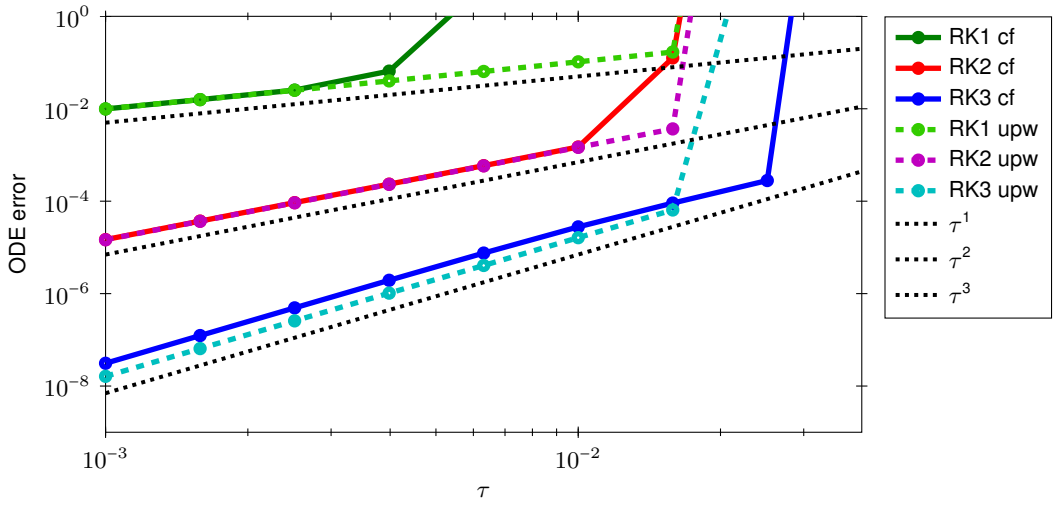
$$\begin{bmatrix} H_{h,x}(t) \\ H_{h,y}(t) \\ E_{h,z}(t) \end{bmatrix} = e^{-t\mathbf{A_{h,TM}}} \begin{bmatrix} H_{h,x}(0) \\ H_{h,y}(0) \\ E_{h,z}(0) \end{bmatrix}.$$

The *ODE error* of the fully discrete approximation is defined as

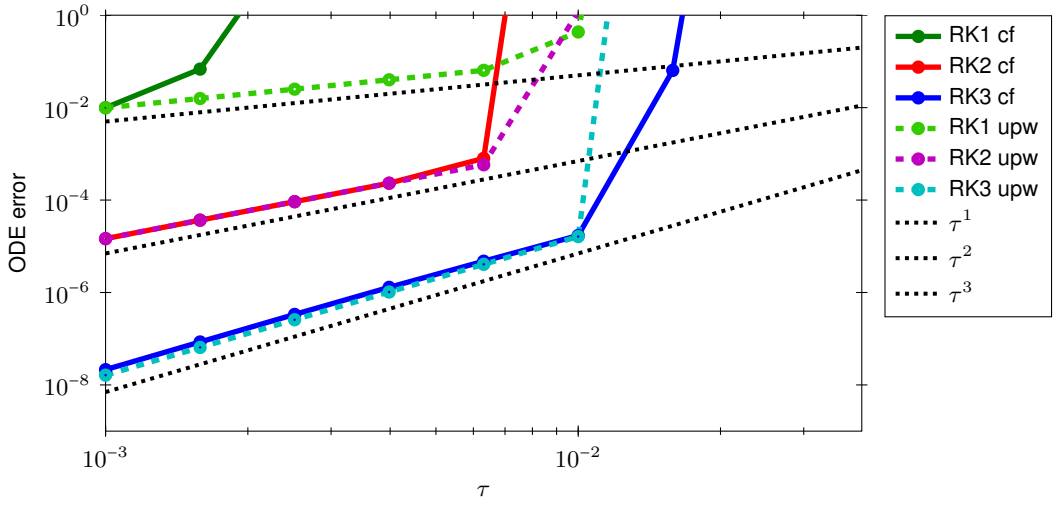
$$e^n_{ODE} := \begin{bmatrix} H^n_{h,x} \\ H^n_{h,y} \\ E^n_{h,z} \end{bmatrix} - \begin{bmatrix} H_{h,x}(t_n) \\ H_{h,y}(t_n) \\ E_{h,z}(t_n) \end{bmatrix}.$$

For this example we use the ODE error to survey the time convergence of the discretizations. Figure 5.4 shows the ODE error for RK1, RK2 and RK3 methods combined with centered or upwind fluxes. We choose a fixed polynomial degree $k = 4$ and use $h \in \{0.5, 0.25, 0.125\}$ as meshsizes. The slope of the plotted errors confirm the proven order τ^s, $s = 1, 2, 3$, for all three RK schemes. In addition, we observe the properties investigated in Section 5.2.1, namely the relaxed CFL conditions of RK1 and RK2 methods when using upwind fluxes instead of centered fluxes and the tightend CFL condition for RK3 in the contrary case.

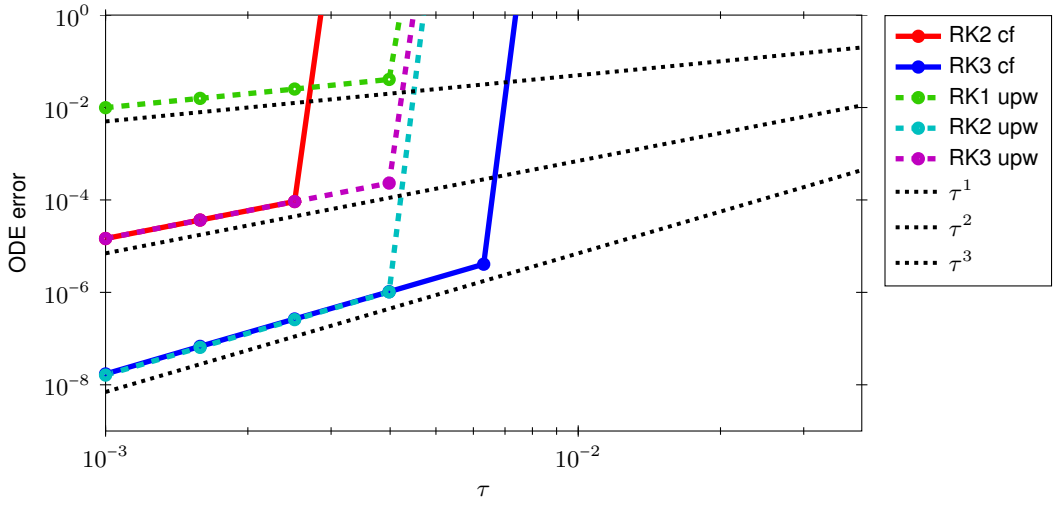
Example 2 Next we consider the example given in Section 5.2.2 to check our results in the inhomogeneous case.We choose the polynomial order $k = 4$ for the spatial discretization. Revisiting the exact solution given in (5.13) we realize that it is also a polynomial of degree 4 in the spatial variables. Consequently, the fully discrete approximation contains no spatial error and the full error e^n is the time discretization error. In Figure 5.5 the full error is plotted for $h \in \{0.5, 0.25, 0.125\}$ and we observe the same behaviour as in the previous example.



(a) meshsize $h = 0.5$

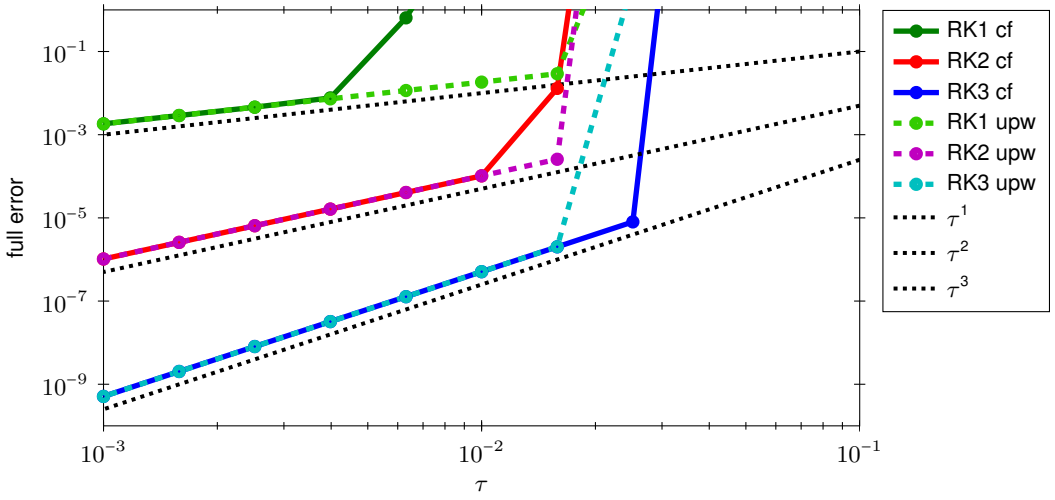


(b) meshsize $h = 0.25$

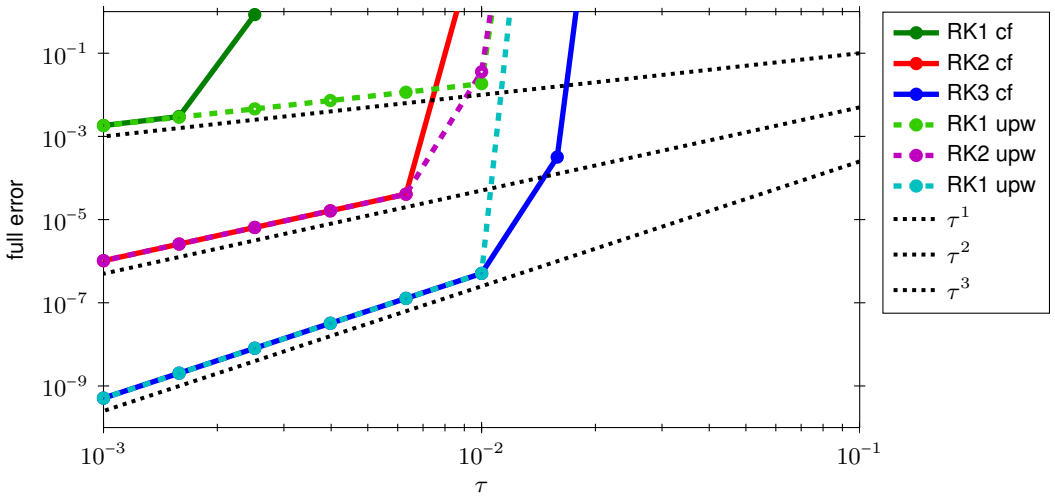


(c) meshsize $h = 0.125$

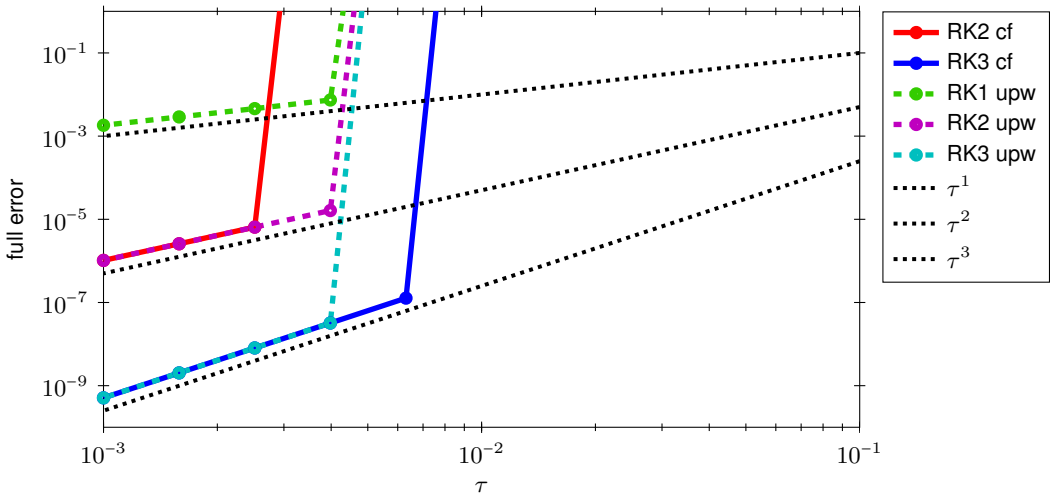
Figure 5.4: Convergence of full discretization



(a) meshsize $h = 0.5$



(b) meshsize $h = 0.25$



(c) meshsize $h = 0.125$

Figure 5.5: Convergence of full discretization

Summary and Outlook

We have provided the stability and the error analysis of dG discretizations of Maxwell's equations in space for both centered and upwind fluxes. Furthermore, we have proven the stability and the convergence of schemes obtained when discretizing the resulting semi-discrete problem with the forward Euler method, RK2 or RK3 methods. Clearly, the next step is to consider RK4 methods, in particular in combination with upwind fluxes. We shortly point out the main problem appearing for the simpler case without source term. Then, we get an energy identity similar to the one of RK3 methods, see (4.25). In fact, we we get

$$\begin{aligned}\|u_h^{n+1}\|_V^2+|u_h^n|_S^2 + \frac{1}{3}\tau|U_h^{n1}|_S^2 + \frac{1}{6}\tau|U_h^{n2}|_S^2 + \frac{1}{2}|U_h^{n3}|_S^2 + \frac{1}{72}\|\tau^3(A_h^{\mathrm{upw}})^3\|_V^2 \\ =\|u_h^n\|_V^2 + \frac{1}{4}\tau|\tau A_h^{\mathrm{upw}}u_h^n|_S^2 + \frac{1}{12}\tau|\tau A_h^{\mathrm{upw}}U_h^{n1}|_S^2 + \frac{1}{24^2}\|\tau^4(A_h^{\mathrm{upw}})^4u_h^n\|_V^2,\end{aligned}$$

with $U_h^{n3} := U_h^{n2} - \frac{1}{6}\tau^3(A_h^{\mathrm{upw}})^3u_h^n$. Even in this case it is not clear how the two S-seminorm terms on the RHS side can be balanced with the terms on the LHS. Since this is the first step towards the convergence our next goal is to solve this problem.

Appendix A

Auxiliary Results

A.1 Stone's Theorem

In Chapter 1 we use Stone's theorem which is stated below.

Theorem A.1 *(Stone's theorem, [18, Theorem 1.36]). Let H be a Hilbert space and $A : \mathcal{D}(A) \to H$ be a linear operator with dense domain, i. e. $\overline{\mathcal{D}(A)} = H$. Then, A generates a C_0-group of unitary operators if and only if A is skew-adjoint.*

A.2 Useful Inequalities

Throughout the thesis the following two inequalities are frequently used.

Theorem A.2 *(Cauchy Schwarz inequality, [19, Theorem I.1.10]). Let (Ω, Σ, μ) be a measure space and let $f, g \in L^2(\mu)$. Then, $fg \in L^1(\mu)$ and there holds*

$$\|fg\|_{L^1} \leq \|f\|_{L^2} \|g\|_{L^2}.$$

Note that this also applies to sequences $a, b \in l^2$. Then, $ab \in l^1$ and it holds

$$\|ab\|_{l^1} \leq \|a\|_{l^2} \|b\|_{l^2}.$$

Theorem A.3 *(Young's inequality). Let $x, y \geq 0$ be real numbers. Then, there holds for every $\gamma > 0$*

$$xy \leq \frac{1}{2}\gamma x^2 + \frac{1}{2}\gamma^{-1} y^2.$$

We call this inequality the weighted Young's inequality. The (usual) Young's inequality is obtained by choosing $\gamma = 1$.

A.3 Gronwall Lemmata

In Chapter 3 and 4 the continuous and discrete Gronwall lemma are used.

Theorem A.4 *(Continuous Gronwall lemma, [6, Proposition 2.1]. Let $T \in \mathbb{R}_+ \cup \{\infty\}$, $f, g \in L^\infty(0, T)$ and $c \geq 0$. Furthermore, let g be a monotonically increasing, continuous function and let f satisfy*

$$f(t) \leq g(t) + c \int_0^t f(s)\, ds \qquad \text{a. e. in } [0, T].$$

Then, there holds

$$f(t) \leq e^{ct} g(t).$$

Theorem A.5 *(Discrete Gronwall lemma, [6, Proposition 4.1]). Let* $\{a_n\}_n$, $\{b_n\}_n \subset \mathbb{R}$ *be two sequences,* $c \geq 0$ *and* $\tau > 0$ *be two constants. Let* $\{b_n\}_n$ *be monotonically increasing and let* $\{a_n\}_n$ *satisfy*

$$a_n \leq b_n + c\tau \sum_{m=0}^{n-1} a_m, \qquad n = 1, 2, \ldots,$$

with initial value $a_0 \leq b_0$*. Then, there holds*

$$a_n \leq (1 + c\tau)^n b_n \leq e^{cn\tau} b_n.$$

Bibliography

[1] F. Brezzi, L. D. Marini, and E. Süli. Discontinuous Galerkin methods for first-order hyperbolic problems. *Math. Models Methods Appl. Sci.*, 14(12):1893–1903, 2004.

[2] E. Burman, A. Ern, and A. Fernández. Explicit Runge-Kutta schemes and finite elements with symmetric stabilization for first-order linear PDE systems. *SIAM J. Numer. Anal.*, 48(6):2019–2042, 2010.

[3] B. Cockburn. Discontinuous Galerkin methods. *ZAMM - Z. angew. Math. Mech.*, 83(11):731–754, 2003.

[4] S. Descombes, S. Lanteri, and L. Moya. Locally implicit time integration strategies in a discontinuous Galerkin method for Maxwell's equations. *J. Sci. Comput.*, 56(1):190–218, 2013.

[5] W. Dörfler, A. Lechleiter, M. Plum, G. Schneider, and C. Wieners. *Photonic Crystals: Mathematical Analysis and Numerical Approximation*. Oberwolfach Seminars, 42. Springer Basel, 2011.

[6] E. Emmrich. Discrete versions of Gronwall's lemma and their application to the numerical analysis of parabolic problems. Preprint No. 637, Fachbereich Mathematik, TU Berlin, July 1999.

[7] E. Hairer, S.P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I: Nonstiff Problems*. Solving Ordinary Differential Equations. Springer, 2009.

[8] J. S. Hesthaven and T. Warburton. *Nodal Discontinuous Galerkin Methods*, volume 54. Springer, texts in applied mathematics edition, 2008.

[9] M. Hochbruck. Numerik 1-4, 2010 - 2012. Lecture Notes.

[10] M. Hochbruck, T. Jahnke, and R. Schnaubelt. Convergence of an ADI splitting for Maxwell's equations. Technical report, Karlsruhe Institute for Technology, 2013.

[11] J. D. Jackson. *Klassische Elektrodynamik*. de Gruyter, 2006.

[12] A. Kirsch. Introduction into Maxwell's Equations. Lecture notes, 2012.

[13] D. Levy and E. Tadmor. From semidiscrete to fully discrete stability of Runge-Kutta schemes by the energy method. *SIAM Review*, 40:40–73, 1998.

[14] P. Monk. *Finite element methods for Maxwells equations*. Clarendon Press, 2006.

[15] J. Niegemann. *High-Order Methods for Solving Maxwell's Equations in the Time-Domain*. PhD thesis, Karlsruhe Institute of Technology, 2009.

[16] T. Pažur. *Error analysis of implicit and exponential time integration of linear Maxwell's equations*. PhD thesis, Karlsruhe Instutute of Technology, 2013.

[17] D. A. Di Pietro and A. Ern. *Mathematical Aspects of Discontinuous Galerkin Methods*. Springer, 2012.

[18] R. Schnaubelt. Evolution equations, 2012. Lecture Notes.

[19] D. Werner. *Funktionalanalysis*. Springer Berlin Heidelberg, 6 edition, 2007.

Erklärung

Hiermit versichere ich, dass ich diese Arbeit selbständig verfasst und keine anderen, als die angegebenen Quellen und Hilfsmittel benutzt, die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht und die Satzung des Karlsruher Instituts für Technologie zur Sicherung guter wissenschaftlicher Praxis in der jeweils gültigen Fassung beachtet habe.

Karlsruhe, den 31.01.2014