Error analysis of splitting methods for wave type equations

Zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

von der KIT-Fakultät für Mathematik
des Karlsruher Instituts für Technologie (KIT)
genehmigte

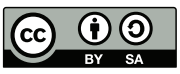
Dissertation

von

Johannes Eilinghoff

Tag der mündlichen Prüfung: 12. Juli 2017

1. Referent: Prof. Dr. Roland Schnaubelt
2. Referentin: Prof. Dr. Marlis Hochbruck



Contents

Introduction

Almost everywhere in science laws of nature are described by ordinary or partial differential equations. Only very rarely and typically only for simple problems the solution can be evaluated directly. For most differential equations and especially for questions coming from applications it is hence necessary to compute solutions numerically.

Over the years many numerical schemes for a large amount of different types of equations have been invented, analysed and tested in practice. Nevertheless, different summands within one equation often do not behave numerically equal and should therefore be treated with methods that are adapted to them. This mainly happens when different phenomena have been included in the differential equation during the mathematical modelling stage.

Splitting methods are a way to tackle these difficulties and to compute in a small amount of time a numerical solution that differs not much from the exact solution. They are well suited to equations where one has an efficient numerical solver for each summand. The basic idea of splitting methods is to combine them to gain a numerical solution of the whole equation by treating the different parts of the equation one after another. The result of each sub-step with one part of the equation is used as initial value for the computation of the next sub-step with another part of the equation. General and detailed information on splitting methods can be found in the survey article [54].

This procedure requires that the problem can be written as an evolution equation that is first-order in time. Then the terms except the one with time derivative determine the rate of change of the observed quantity. At least for small time step sizes it is reasonable to assume that it makes not much difference whether the summands of the rate of change are treated together or one after another. The precise dependency of this difference on the time step size is quantified by the convergence order of the numerical scheme. The most important topic of this thesis is to prove convergence orders of splitting schemes. For the investigation of the topic it is crucial which norm is chosen for the errors estimates.

A further reason for using schemes that treat each part in an appropriate way is that they often conserve the energy, the momentum, the positivity or the regularity of a solution.

If there exist already implemented algorithms for some types of equations, it is fairly easy to combine them to a splitting method. This allows to compute solutions to more involved equations that contain these well-known types of equations without having to write the complete code from the scratch. This gain of programming time is especially

an important advantage in applications.

When dealing with the numerical computation it should not be forgotten to assure that the differential equation has a unique solution (in a suitable sense) since it is useless to compute an approximation to a solution that does not exist and since we can hardly say to which solution the approximation belongs if there is more than one solution.

Although it does not appear in this thesis, we mention that boundary conditions can cause an order reduction of a scheme, sometimes in a rather unexpected way. A remedy to this can be a different splitting of the right-hand side of the equation, see [20] and [21].

A general technical problem in the theoretical analysis of splitting methods is that often a high spatial regularity of the initial functions and the solutions is necessary. As a consequence, the lack of regularity can reduce the convergence order of the scheme, see Chapter 4 and 5, as well as Section 10.3.

In the thesis at hand we tackle two partial differential equations from physics with different types of splitting schemes: the cubic nonlinear Schrödinger equation with exponential splitting methods and the Maxwell equations with an ADI splitting method. It might be possible to treat other wave type equations, like the nonlinear wave equation, with similar techniques as the ones presented in this thesis.

In practical computations always space discretization errors come into play. In this thesis we restrict ourselves to the time discretization errors and do not give an error analysis of the full discretizations.

Exponential splitting methods for nonlinear Schrödinger equations

We analyse the convergence order of two splitting schemes applied to the cubic nonlinear Schrödinger equation on the torus and on the full space. The linear part is treated with the fast Fourier transform and for the solution of the nonlinear part we use the existing explicit solution formula. We start with a well-known theorem by C. Lubich from [51] and put our main focus on the question whether (and to what extend) a reduction of the regularity of the initial function causes a reduction of the convergence order. This turns out to be true and can be found together with the proof in part two of this thesis. We add numerical experiments to investigate the order reduction in practice. We have published the theoretical analysis in [22].

The earlier paper [10] contains a convergence result for the case of two space dimensions and any globally Lipschitz nonlinearity. Defect-based local error estimators for the nonlinear Schrödinger equation were proven in [7] (see also [5] and [6] for the linear case). Adaptive splitting methods for the Schrödinger equation in the semiclassical regime were studied in [4]. An analysis of the cubic semilinear Schrödinger equation with damping and forcing terms on the torus for regular initial functions can be found in [44]. Low regularity exponential-type integrators for the cubic nonlinear Schrödinger equation were investigated in [60] very recently.

The long-time behaviour of numerical (splitting) schemes for a spectral semi-discretization of nonlinear Schrödinger equations was investigated in [28] and [29], see also [26] and [25]. For a quasilinear Schrödinger equation and solutions in H^7, the paper [50] provides error estimates in H^1 of the Strang splitting combined with a frequency cut-off.

In contrast to [25] or [51], we do not use Lie derivatives and Lie commutators to show the local error estimates. Instead we employ error formulas that are derived by iterating the solution formula and by replacing the exponential function in the numerical scheme by a Taylor expansion, see [12] for a similar procedure. We split the error formulas into a quadrature error and several remainder terms as in e.g. [12], [43] and [27]. The main novelty of our approach is the use of fractional convergence results. They allow us to treat initial values in spaces larger than H^4 (which was taken in [51]). Moreover, for the Lie splitting the fractional convergence in $H^{7/4}$ is crucial for the necessary a priori bound in $H^{7/4}$ of the numerical solution. The needed estimates, involving fractional orders of the time step size, are established by various interpolation arguments, e.g. when controlling quadrature errors.

An ADI method for the Maxwell equations

The other problem from physics we address are the Maxwell equations. For them we use an alternating direction implicit (ADI) scheme that is based on the splitting of the curl operator into those partial derivatives with negative and those with positive signs in the Maxwell operator. We deal with the error estimate and the convergence order of the ADI method and add an analysis of the preservation of the divergence identity. The main advantage of the ADI method we investigate is its efficiency. We can rewrite the resulting equations in such a way that systems of three-dimensional implicit equations decouple into three one-dimensional implicit equations. We conclude that part of this thesis with numerical experiments that confirm some of our results.

The idea of ADI methods in general was published in [63] for the heat equation. The studies therein were further developed in [18] and [19]. An analysis of dimension splitting methods for abstract evolution equations was done in [35].

We compute the space derivatives with finite differences on the Yee grid, as proposed in [72]. This combination was first done in [76] and [75]. An analysis of the numerical dispersion was done in [74] and a combination with perfectly matched layers was investigated in [49] and [30]. A version of this scheme for the two-dimensional Maxwell equations was discussed in [57]. A much earlier approach of a combination of an ADI scheme with the Yee grid was presented in [41].

The ADI splitting we present is not the only possible one, see for instance [14]. Finite element methods with an explicit time integration scheme on a spatial mesh that contains very small mesh elements often come along with severe CFL conditions on the time step size. This difficulty can be overcome by an implicit method. An approach to the Maxwell

equations with a locally implicit method to avoid this difficulty was investigated in [39].

In our splitting method we use resolvents of splitting operators, so that it belongs to the class of resolvent splitting methods. An abstract analysis of two different resolvent splitting methods was done in [59].

Structure of this thesis

This PhD thesis consists of three parts and is organised as follows.

The first part is Chapter 1 and contains an overview over splitting methods in Section 1.3. Some notations and preliminaries are denoted in Section 1.1, while Section 1.2 gives an introduction into quadrature rules. Important theorems from functional analysis and semigroup theory that we use in this thesis are recalled in Section 1.4.

In the second part of this thesis we deal with splitting methods for nonlinear Schrödinger equations. In Chapter 2 we state the problem we are looking at for the rest of this part. Section 2.1 contains well-known facts about nonlinear Schrödinger equations, especially on the well-posedness theory, and we introduce the splitting schemes we use. The functional analytic setting for our analysis is presented in Section 2.2. From then on we restrict ourselves to the case of a cubic nonlinearity.

Chapter 3 is devoted to the situation that the initial function is in H^4. This situation was already investigated for the case of the torus in [51]. In Section 3.1 we state that the Strang splitting scheme converges in L^2 with order two in the time step size to the exact solution. We additionally note auxiliary results that appear in the proof of this theorem. This proof consists of arguments in H^2, followed by considerations in L^2. They are presented in detail in the Sections 3.2 and 3.3, respectively.

Our main contribution to the scientific progress from this part is the convergence theorem for initial functions in $H^{2+2\theta}$ for $\theta \in (0,1)$. It reads that the convergence order in L^2 reduces to $1+\theta$ and is the topic of Chapter 4. We present the theorem itself as well as intermediate results in Section 4.1. The proof follows the same structure as the one for the theorem in the H^4-situation in Chapter 3 and is the content of the Sections 4.2 and 4.3.

Finally, we investigate the situation that the initial function has only H^2-regularity. We are able to show in Chapter 5 that in this case the Lie and the Strang splitting are convergent of order one in L^2. As far as we know it is the first result in this setting for the Lie splitting. Section 5.1 contains these theorems and the most important results required in their proofs. In Section 5.2 we show the proofs of the statements.

We close part two of this thesis by numerical experiments in Chapter 6. We conduct them to confirm the theoretical results we have shown in the previous chapters and to show their sharpness. After giving an overview over the experiments in Section 6.1, we explain in Section 6.2 the two techniques we use to gain initial functions of a given regularity. We test our code in Section 6.3 on the example of plane wave solutions, for which the formula

of the solution is known explicitly, and on the example of modified soliton solutions. In Section 6.4 we compute the numerical convergence order of the scheme for initial functions belonging to several H^s-spaces and see the reduction of the convergence order we have shown in Chapter 4. In Section 6.5 we see in an experiment that the error constant increases for highly oscillating initial functions.

In the third part of the thesis we analyse an alternating direction implicit (ADI) splitting for the Maxwell equations. In Chapter 7 we describe the problem we look at and show properties of its solutions. We introduce the Maxwell equations in Section 7.1. The functional analytic setting for this part and the introduction of the Maxwell operator and the splitting operators, as well as the proofs of basic properties of them and some embedding theorems, are contained in Section 7.2. In Section 7.3 we prove the well-posedness of the problem and additionally embedding and trace properties of the domain of the Maxwell operator and the three restrictions of the Maxwell operator we use.

Chapter 8 is devoted to the properties of the splitting operators and the ADI splitting scheme. In the Sections 8.1, 8.2 and 8.3 we show that the splitting operators generate C_0-semigroups and that their resolvents satisfy some estimates. We introduce the ADI scheme we work with in Section 8.4 and close the chapter with a proof of its efficiency in Section 8.5. This efficiency is the main reason for using the ADI scheme.

In Chapter 9 we use the properties of the splitting operators that have been shown in Chapter 8 to prove the convergence of the scheme. In Section 9.1 we show the convergence of order one in L^2 and in Section 9.2 we use similar techniques to prove the same result in a weak sense. The exact solution of the Maxwell equations satisfies two identities involving the divergence of the electric and the magnetic field, respectively. These equations are satisfied by the numerical solutions in a weak sense and in L^2, which we see in the Sections 9.3 and 9.4, respectively.

The last chapter of this thesis, Chapter 10, is devoted to the numerical verification of the theoretical results in L^2 of the ADI scheme. In Section 10.1 we give an overview over our numerical experiments. We conduct two experiments in Section 10.2 for the situation without conductivity and without external currents, in which the exact solution is known. The results help us not only to check our programming code but also to estimate the appropriate fineness of the space discretization. Afterwards we see the predicted convergence order of the method and the predicted order of the preservation of the divergence properties for the situation with conductivity and external current. We close this chapter by an experiment in Section 10.3 that shows the behaviour of the scheme for the case that the initial function does not satisfy all regularity assumptions of the convergence theorem in Section 9.1.

Acknowledgements:

First, I thank my supervisor Prof. Dr. Roland Schnaubelt for giving me this interesting

topic for my PhD thesis. We had many fruitful discussions about my research and he gave me a lot of suggestions and helpful advice how to continue and improve my work. I thank my second supervisor Prof. Dr. Marlis Hochbruck for many helpful comments on my research, mainly on the programming of the numerical examples. Concerning an effective use of the MATLAB software for my simulations I appreciate Prof. Dr. Tobias Jahnke's help. I thank JProf. Dr. Katharina Schratz for productive discussions during our joint work on the cubic nonlinear Schrödinger equation. I thank my colleagues of the workgroup "Functional analysis" at the Institute for Analysis and of other workgroups at the KIT Department of Mathematics for a great working environment and a pleasant collegiality.

I am very grateful that my supervisors gave me the opportunity to take part in many excellent workshops and conferences, to listen to interesting talks and lectures, and to meet fellow scientists from all over the world. I thank all the speakers and lecturers I had the pleasure to learn from during the last ten years (which include my diploma studies). They made it possible for me to gather enough mathematical knowledge and skills to write this PhD thesis. Moreover, they encouraged me with their enthusiasm to tackle unsolved mathematical problems.

I appreciate that apart from meetings on mathematical research I was allowed to attend English courses as well as the KIT mentoring programme X-Ment, and to obtain the "Hochschuldidaktikzertifikat Baden-Württemberg". They are a help in my daily working life and prepared me for business tasks that may come in future.

Besides my research, I gave exercise classes for first-year students of the major subjects mathematics and civil engineering. This gave me the opportunity to refine my teaching skills and to pass over my knowledge and enthusiasm for mathematics to the next generation of students.

I thank many of my colleagues, amongst them Sven Caspart, Andreas Geyer-Schulz, Christine Grathwohl, Fabian Hornung, Luca Hornung, Marcel Mikl, Tobias Ried, Sebastian Schwarz and Andreas Sturm, for discussions about my research or for reading parts of this thesis.

Working on a PhD thesis is not possible without having funding. I thank the State of Baden-Württemberg, the KIT Department of Mathematics, and the Deutsche Forschungsgemeinschaft in the framework of the Research Training Group 1294 "Analysis, Simulation and Design of Nanotechnological Processes" and the Collaborative Research Center 1173 "Wave phenomena: analysis and numerics", for financially supporting my research.

Last but not least I thank my family and my friends for encouraging and supporting me during the last years.

Johannes Eilinghoff
July 2017

Part I.

Splitting methods in general

1. Introduction to splitting methods

In this chapter we present the basic principles of splitting methods in general and explain mathematical background needed later on. We start with notations and concepts from functional analysis in Section 1.1. Afterwards Section 1.2 gives an overview over quadrature rules. Splitting methods are motivated and explained in Section 1.3. We present the two types of them we use and comment on their basic properties. The chapter is closed by Section 1.4 with a collection of important theorems from functional analysis and semigroup theory used in this thesis.

1.1. Notations and preliminaries

Throughout this thesis c denotes a generic constant, whose values may change from appearance to appearance, also within the same equation. It possibly depend on the dimension of the spatial set on which our differential equations are defined and on embedding constants. Moreover, I is the identity operator, $\mathbb{1}$ the function being constant one and $\mathbb{1}_A$ the indicator function of a set A, i.e. $\mathbb{1}_A(x) = 1$ if $x \in A$ and 0 otherwise.

Let X and Y be two Banach spaces. We write $Y \hookrightarrow X$ if Y is continuously embedded into X and $X \cong Y$ if there exists an isomorphism between X and Y. We denote the duality pairing of Y^* and Y by $\langle y^*, y\rangle_{Y^*,Y}$ or by $\langle y, y^*\rangle_{Y,Y^*}$ for $y \in Y$ and $y^* \in Y^*$. If X is a Hilbert space, we write $(\cdot \mid \cdot)_X$ for its inner product. Note that if Y is densely embedded into X and if X is a Hilbert space, we have $\langle x, y\rangle_{Y^*,Y} = (x \mid y)_X$ for $x \in X \cong X^* \hookrightarrow Y^*$ and $y \in Y \hookrightarrow X$.

The Banach space of all bounded linear operators from X to Y is denoted by $\mathcal{B}(X, Y)$, and by $\mathcal{B}(X)$ if $Y = X$. The domain $D(A)$ of a linear operator $A : D(A) \subseteq X \to X$ is always equipped with the graph norm, which is defined by $\|x\|_{D(A)} := \|x\|_X + \|Ax\|_X$ for $x \in D(A)$. The resolvent of such a linear operator is denoted by $(\lambda I - A)^{-1}$ for λ being in the resolvent set of A. Linear operators act on all expressions that follow till the enclosing parenthesis end or till the summand in which they appear ends. The *part of a linear operator* $A : D(A) \subseteq X \to X$ *in a subspace* $Y \subseteq X$ is the operator $A_Y : D(A_Y) \subseteq Y \to Y$ with

$$D(A_Y) := \{y \in Y \mid y \in D(A), Ay \in Y\}$$

and $A_Y y = Ay$ for all $y \in D(A_Y)$.

Let $\Omega \subseteq \mathbb{R}^d$ be an open set with the spatial dimension $d \in \mathbb{N}$. The set of infinitely often differentiable real-valued or complex-valued functions with compact support in Ω is denoted by $C_c^\infty(\Omega)$. Let $p \in [1, \infty]$ and $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$. The *Lebesgue spaces* are the Banach spaces defined by

$$L^p(\Omega) := \Big\{ f : \Omega \to \mathbb{K} \text{ Borel measurable } \Big| \int_\Omega |f(x)|^p \, \mathrm{d}x < \infty \Big\}, \qquad p \in [1, \infty),$$
$$L^\infty(\Omega) := \{ f : \Omega \to \mathbb{K} \text{ Borel measurable } | \, \exists c \geq 0 : |f(x)| \leq c \text{ for almost all } x \in \Omega \}, \qquad p = \infty,$$

and are equipped them with the norms

$$\|f\|_{L^p} := \left(\int_\Omega |f(x)|^p \, \mathrm{d}x \right)^{1/p}, \qquad p \in [1, \infty),$$
$$\|f\|_\infty := \|f\|_{L^\infty} := \inf_{c \geq 0} \big\{ |f(x)| \leq c \text{ for almost all } x \in \Omega \}, \qquad p = \infty.$$

In the same way we define the Lebesgue spaces for non-open Borel measurable sets $\Omega \subseteq \mathbb{R}^d$.

Furthermore, we define for a non-empty open set $\Omega \subseteq \mathbb{R}^d$ the weak derivatives and the Sobolev spaces. We denote by $L^1_{loc}(\Omega)$ the space of all Borel measurable *locally integrable functions*, i.e. all Borel measurable $f : \Omega \to \mathbb{K}$ for which the restriction $f|_K$ to any compact set $K \subseteq \Omega$ is in $L^1(K)$. Let $f \in L^1_{loc}(\Omega)$. It is weakly differentiable with respect to the j-th variable if there exists a $g \in L^1_{loc}(\Omega)$ such that

$$\int_\Omega f \partial_j \varphi \, \mathrm{d}x = - \int_\Omega g \varphi \, \mathrm{d}x$$

for all $\varphi \in C_c^\infty(\Omega)$. In this case g is called *weak derivative* of f and we write $\partial_j f$ for g. Weak derivatives of higher order are defined recursively. The order of a multiindex $\alpha \in \mathbb{N}^d$ is defined by $\alpha_1 + \cdots + \alpha_d$ and denoted by $|\alpha|$. Observing that weak derivatives commute, we denote a weak derivative with respect to α as $\partial^\alpha := \partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d}$. For $k \in \mathbb{N}_0$ and $p \in [1, \infty]$ we introduce the *Sobolov space* of order k as

$$W^{k,p}(\Omega) := \big\{ f \in L^p(\Omega) \, \big| \, \partial^\alpha f \text{ exists and } \partial^\alpha f \in L^p(\Omega) \text{ for all } \alpha \in \mathbb{N}^d \text{ with } |\alpha| \leq k \big\}$$

and equip it with the norm

$$\|f\|_{W^{k,p}} := \left(\sum_{\alpha \in \mathbb{N}^d, \ |\alpha| \leq k} \|\partial^\alpha f\|_{L^p}^p \right)^{1/p}, \qquad p \in [1, \infty),$$
$$\|f\|_{W^{k,\infty}} := \max_{\alpha \in \mathbb{N}^d, \ |\alpha| \leq k} \|\partial^\alpha f\|_{L^\infty}, \qquad p = \infty.$$

With real interpolation theory, see Section 7.57 in [1], we define the *fractional Sobolev spaces* for $s \geq 0$, $k \in \mathbb{N}_0$ with $s \leq k$, and $p \in [1, \infty]$ as

$$W^{s,p}(\Omega) := \big(L^p(\Omega), W^{k,p}(\Omega) \big)_{s/k,p},$$

equipped with the norm given by the interpolation. All Sobolev spaces are Banach spaces. In the case $p = 2$, which is the most important case for this thesis, they are Hilbert spaces and we write $H^s(\Omega) := W^{s,2}(\Omega)$. Note that $H^0(\Omega) = L^2(\Omega)$. We use real and also complex interpolation of Hilbert spaces in this thesis. For further information about these topics we refer to [52]. Sobolev spaces are discussed in detail in [1].

If $f \in W^{1,\infty}(\Omega) \cap W^{2,3}(\Omega)$, then we define the norm

$$\|f\|_{W^{1,\infty}\cap W^{2,3}} := \|f\|_{W^{1,\infty}} + \|f\|_{W^{2,3}} .$$

Let $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the unitary *Fourier transform* and its unitary inverse on $L^2(\mathbb{R}^d)$ and on $L^2(\mathbb{T}^d)$, respectively. For $\Omega \in \{\mathbb{R}^d, \mathbb{T}^d\}$ and all $s \geq 0$ there exists the characterization

$$H^s(\Omega) = \left\{ f \in L^2(\Omega) \mid \mathcal{F}^{-1}\left((1 + |\xi|^2)^{s/2}\mathcal{F}f\right) \in L^2(\Omega) \right\} \tag{1.1}$$

for the Sobolev spaces $H^s(\Omega)$ and

$$\|f\|_{H^s} \simeq \left\|\mathcal{F}^{-1}\left((1 + |\xi|^2)^{s/2}\mathcal{F}f\right)\right\|_{L^2} = \left\|(1 + |\xi|^2)^{s/2}\mathcal{F}f\right\|_{L^2} ,$$

for their norms, see Section 7.62 in [1] for the case $\Omega = \mathbb{R}^n$. Thereby, $\simeq$ means equal up to a multiplicative constant. We remark that for $s \in [0, 4]$ this norm equivalence holds true with constants independent of s by taking $k = 4$ in the definition of the fractional Sobolev spaces and interpolating between the norm estimates in L^2 and H^k. On the torus we actually have the norm in $\ell^2(\mathbb{Z}^d)$ on the right-hand side of the above identity. As above, we suppress the domain in the notation of norms if the main spatial domain in the corresponding context is meant.

1.2. Quadrature rules

In numerical analysis it is often necessary to compute the value of an integral over a continuous function f. We need that in this thesis to estimate differences of an integral and evaluations of functions appearing in Taylor expansions, and to incorporate inhomogeneities into numerical schemes. If it is not possible to calculate the exact value of the integral, we have to approximate it numerically. This can be done with *quadrature rules*.

We first consider one-dimensional integrals. Let H be a Hilbert space with norm $\|\cdot\|_H$ and let $f \in C([0, 1], H)$ be a function. Looking at $\int_0^1 f(t)\,\mathrm{d}t$, we evaluate f at certain points in the interval $[0, 1]$ and sums these function values up after multiplying them with certain weights. So, a quadrature rule is given by a number $n \in \mathbb{N}$, *nodes* $0 \leq c_1 < \cdots < c_n \leq 1$ and *weights* $\omega_i \geq 0$, $i = 1, \ldots, n$. It approximates the integral by

$$\int_0^1 f(t)\,\mathrm{d}t \approx \sum_{i=1}^{n} \omega_i f(c_i).$$

We pose the restriction $\sum_{i=1}^{n} \omega_i = 1$ since at least constant functions shall be integrated without error.

All quadrature rules can be carried over to other intervals via translations and dilations. On the interval $[t_0, t_0 + \tau]$, which is the case we mostly need, they read as

$$\int_{t_0}^{t_0+\tau} f(t)\,\mathrm{d}t \approx \tau \sum_{i=1}^{n} \omega_i f(t_0 + c_i\tau).$$

A quadrature rule is said to be of *order* $k \in \mathbb{N}$ if every polynomial with degree at most $k-1$ is integrated exactly. It is easy to see that this is the case if and only if for all $l = 1, \ldots, k$ we have

$$\sum_{j=1}^{n} \omega_j c_j^{l-1} = \frac{1}{l}.$$

The following error estimate is well-known and can for instance be found as Theorem 3.2.2 in [65]. Its scalar-valued proof transfers directly to the Hilbert space-valued situation.

Proposition 1.1. *Let a quadrature rule be given by $n \in \mathbb{N}$, nodes $0 \le c_1 < \cdots < c_n \le 1$ and weights $\omega_i \ge 0$ for all $i \in \{1, \ldots, n\}$ that has (at least) order k. Let f be (at least) k-times continuously differentiable on $[t_0, t_0 + \tau]$. Then we have the error estimate*

$$\left\| \int_{t_0}^{t_0+\tau} f(t)\,\mathrm{d}t - \tau \sum_{i=1}^{n} \omega_i f(t_0 + c_i\tau) \right\|_H \le c\tau^{k+1} \max_{s\in[t_0,t_0+\tau]} \left\| f^{(k)}(s) \right\|_H.$$

The simplest quadrature rule is the *rectangular rule*. More precisely, there is the rectangular rule with the left endpoint and the rectangular rule with the right endpoint. They have the single node $c_1 = 0$ or $c_1 = 1$, respectively, and the weight $\omega_1 = 1$, so that

$$\int_{t_0}^{t_0+\tau} f(t)\,\mathrm{d}t \approx \tau f(t_0) \qquad \text{and} \qquad \int_{t_0}^{t_0+\tau} f(t)\,\mathrm{d}t \approx \tau f(t_0 + \tau),$$

respectively. They are both of order one.

The *midpoint rule* also has only one node, $c_1 = 1/2$, and one weight, $\omega_1 = 1$, but is of order two. A quadrature rule with the same order is the *trapezoidal rule*, which has the two nodes $c_1 = 0$ and $c_2 = 1$ and the weights $\omega_1 = \omega_2 = 1/2$. In formulas these two rules read

$$\int_{t_0}^{t_0+\tau} f(t)\,\mathrm{d}t \approx \tau f(t_0 + \tau/2) \qquad \text{and} \qquad \int_{t_0}^{t_0+\tau} f(t)\,\mathrm{d}t \approx \frac{\tau}{2}\big(f(t_0) + f(t_0 + \tau)\big).$$

We further mention the second order quadrature rule with the three nodes $c_1 = 0$, $c_2 = 1/2$ and $c_2 = 1$ and the weights $\omega_1 = \omega_3 = 1/4$ and $\omega_2 = 1/2$, i.e.

$$\int_{t_0}^{t_0+\tau} f(t)\,\mathrm{d}t \approx \frac{\tau}{4} f(t_0) + \frac{\tau}{2} f(t_0 + \tau/2) + \frac{\tau}{4} f(t_0 + \tau),$$

which appears in the Sections 9.3 and 9.4. Observe that there exists an order-four quadrature rule with the same nodes, namely the *Simpson rule* with the weights $\omega_1 = \omega_3 = 1/6$ and $\omega_2 = 2/3$, reading

$$\int_{t_0}^{t_0+\tau} f(t)\,\mathrm{d}t \approx \frac{\tau}{6} f(t_0) + \frac{2\tau}{3} f(t_0 + \tau/2) + \frac{\tau}{6} f(t_0 + \tau),$$

Unfortunately, we cannot use the Simpson rule in the above mentioned sections since the weights $\omega_1 = \omega_3 = 1/4$ and $\omega_2 = 1/2$ come out of the proof of the error formulas.

We can also define multidimensional quadrature rules, which we do in this thesis with a two-dimensional rule that approximates an integral over a simplex. The standard two-dimensional simplex is the set

$$S := \{(x, y) \in \mathbb{R}^2 \mid x, y \geq 0, x + y \leq 1\}.$$

For a function $f \in C(S, H)$ we use the approximation

$$\int_S f(x, y)\,\mathrm{d}(x, y) \approx \frac{1}{8}\big(f(0, 0) + f(1, 0) + f(0, 1) + f(1/3, 1/3)\big).$$

We will see in Lemma 3.9 that this quadrature rule is of order two.

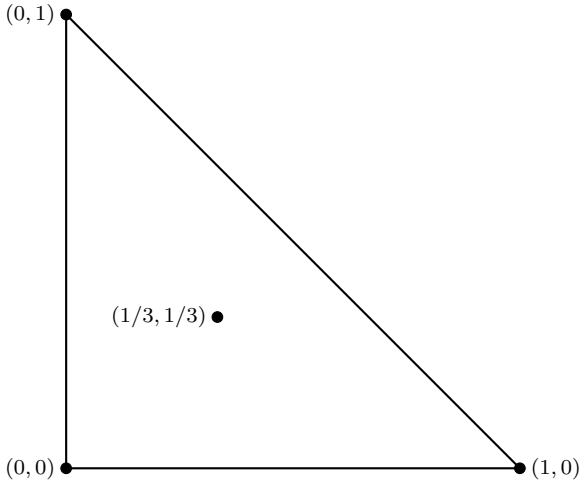


Figure 1.1.: Two-dimensional simplex with the nodes of the quadrature rule.

More information on quadrature rules can be found for example in Section 3 in [65].

1.3. Splitting methods

1.3.1. The idea of splitting methods

We consider a differential equation of the type

$$u'(t) = Lu(t)$$

together with an initial time t_0 and an initial condition $u(t_0) = u_0$. Assume that the operator L can be written as the sum of two operators A and B, i.e. we consider the problem

$$\begin{cases} u'(t) = Au(t) + Bu(t), \quad t \geq t_0, \\ u(t_0) = u_0. \end{cases} \tag{1.2}$$

We suppose that problem (1.2) has a unique solution on the time interval $[t_0, T]$ for a $T > t_0$. Our goal is to compute numerically an approximate solution to problem (1.2) with as little amount of computation costs as possible.

Therefore, we look at the two "subproblems"

$$\begin{cases} v'(t) = Av(t), \quad t \geq t_0, \\ v(t_0) = v_0, \end{cases} \tag{1.3a}$$

and

$$\begin{cases} w'(t) = Bw(t), \quad t \geq t_0, \\ w(t_0) = w_0. \end{cases} \tag{1.3b}$$

We assume that they both have a unique solution on the time interval $[t_0, T]$ and that these solutions can be computed efficiently. Thus, a computer needs only a small amount of time for computing an approximate solution that differs not much from the exact solution. Examples for operators for which the corresponding problem can be solved efficiently are the cases when the solutions of (1.3) are explicitly given or have a simple representation in the Fourier mode, as the Laplace operator on the torus for instance. This is precisely the situation we have in the second part of this thesis for the nonlinear Schrödinger equation.

The idea of *splitting methods* is to exploit the good solvability properties of (1.3) to get an approximate solution for (1.2) in the following way (based on the Lie splitting, see below). We fix a time step size $\tau = \frac{T-t_0}{N} > 0$ for an $N \in \mathbb{N}$ and calculate the solution v of the first subproblem with initial function u_0 after one time step of length τ. Then we define $\widetilde{u}_1 := v(t_0+\tau) = e^{tA}u_0$ and calculate the solution w of the second subproblem with initial function $\widetilde{u}_1$ (and again starting time t_0), getting $u_1 := w(t_0+\tau) = e^{tB}\widetilde{u}_1 = e^{tB}e^{tA}u_0$. The function u_1 is now taken as the approximate solution of problem (1.2) at time $t_0+\tau$. Afterwards we repeat this procedure with initial function u_1 as initial function for the first subproblem until we reach the end time T of our computation. A graphical illustration of this approach is displayed on the left-hand side of Figure 1.2.

The described procedure causes as *time discretization error* the so-called *splitting error*, which is due to the fact that we only compute solutions of the subproblems (1.3) and never of the original problem (1.2). Fortunately, there is hope that for small time step sizes the error is small. The reason for this optimism is that (1.2) is a differential equations of first order in time, which means that the right-hand side depicts the rate of change of the

solution. For small time step sizes it is plausible that it does not make a huge difference whether we treat both summands of the rate of change at once or one after another.

One of the most important questions concerning splitting methods is the one for their *convergence order*. The convergence order is the rate with which the time discretization error of the approximation decreases when the time step size is reduced. The most important topic in this thesis is to determine convergence orders of splitting schemes.

The idea of splitting methods can be generalized in a straightforward way to a sum of finitely many operators $L := A_1 + \cdots + A_m$, $m \in \mathbb{N}$. With the help of quadrature rules it is also possible to include inhomogeneities, see for example Subsection 1.3.3.

In this thesis we deal with two types of splitting methods. We use exponential splitting methods in Chapter 2 till 6 to tackle the nonlinear Schrödinger equation, while we investigate in Chapter 7 till 10 an application of an alternating direction implicit (ADI) method to the Maxwell equations. Further splitting methods and an overview over splitting methods in general can be found in the survey article [54].

1.3.2. Exponential splitting methods

A relatively obvious type of splitting methods are the exponential splitting methods. They mimic closely the general idea of splitting methods we described in Subsection 1.3.1. A convergence analysis of exponential splitting schemes in an abstract framework was performed in [36]. General information on exponential integrators can be found in the survey article [38].

An *exponential splitting method* is defined by a time step size $\tau > 0$, an $l \in \mathbb{N}$ and coefficients $a_1, \ldots, a_l, b_1, \ldots, b_l \in \mathbb{R}$. In this thesis we only consider methods with the condition $\sum_{k=1}^{l} a_k = \sum_{k=1}^{l} b_k = 1$. This means that we proceed per application of the scheme in total exactly one time step along the solutions of the both subproblems.

Denoting the exact solutions of the subproblems (1.3) by $e^{tA}v_0$ and $e^{tB}w_0$, the result of an exponential splitting method after one time step reads

$$u_1 = e^{b_l \tau B} e^{a_l \tau A} \cdots e^{b_1 \tau B} e^{a_1 \tau A} u_0. \tag{1.4}$$

The three most important exponential splitting methods are the following ones.

The method we described as motivation in Subsection 1.3.1 is the *Lie splitting* (sometimes also called Lie–Trotter splitting), cf. [71]. It is defined by $n = 1$ and $a_1 = b_1 = 1$, i.e. the numerical solution after one time step is given by

$$u_1 = e^{\tau B} e^{\tau A} u_0.$$

The second simplest method is the *Strang splitting* (sometimes also called Strang–Marchuk splitting), which is given by $l = 2$, $a_1 = a_2 = \frac{1}{2}$, $b_1 = 1$ and $b_2 = 0$. It has been

introduced independently in [66] and in [53]. Its result after one time step is given by

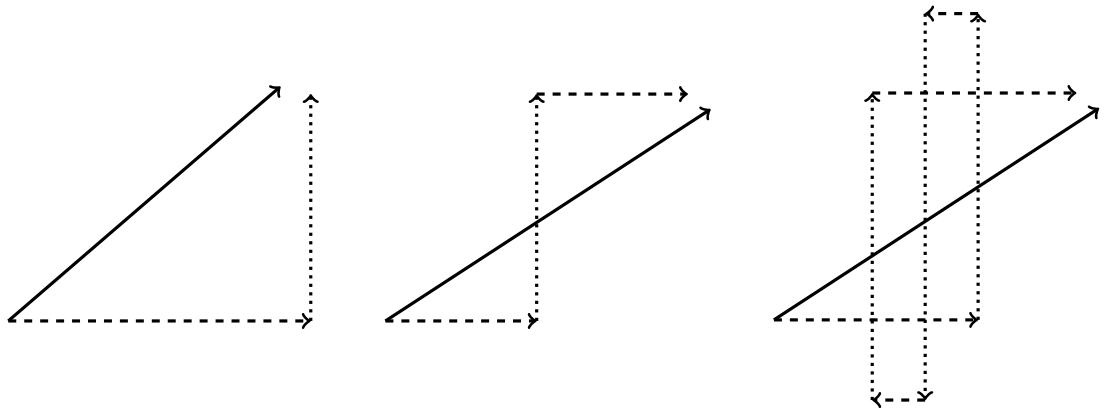
$$u_1 = e^{\frac{1}{2}\tau A} e^{\tau B} e^{\frac{1}{2}\tau A} u_0.$$

The last exponential splitting method we mention is the *Yoshida splitting*, see [73]. It is given by $l = 4$ and the coefficients

$$a_1 = a_4 = \frac{1}{2\big(2 - 2^{1/3}\big)}, \qquad a_2 = a_3 = \frac{1 - 2^{1/3}}{2\big(2 - 2^{1/3}\big)},$$
$$b_1 = b_3 = \frac{1}{2 - 2^{1/3}}, \qquad b_2 = -\frac{2^{1/3}}{2 - 2^{1/3}} \qquad \text{and} \qquad b_4 = 0.$$

The Yoshida splitting has the disadvantage that it uses negative time steps, represented by the arrows going leftwards or downwards in Figure 1.2. This is not an obstacle for hyperbolic problems such as wave type equations due to their time reverseness. But the Yoshida splitting should not be used for parabolic problems since they are not well-defined for negative times. We remark that the Yoshida splitting can also be obtained by composing the Strang splitting with itself ("triple jump method"), see Chapter II in [33].

The Lie, the Strang and the Yoshida splitting are schematically sketched in Figure 1.2.



(a) Lie splitting scheme (b) Strang splitting scheme (c) Yoshida splitting scheme

Figure 1.2.: Schematical sketches of the Lie, the Strang and the Yoshida splitting. The two solutions referring to the operators A and B are drawn as dashed arrows in the horizontal and dotted arrows in the vertical direction, respectively. The solid lines represent the solution of the original problem, having a slightly different end point than the numerical schemes.

Each splitting method has a *classical order*. It is obtained by making a formal Taylor expansion of (1.4) and comparing the terms with a Taylor expansion of the exact solution $e^{t(A+B)}u_0$. Regardless of the given problem and the regularity of the initial function, the

order of a splitting method can never be higher than its classical order. For the three splitting methods introduced above we have the following classical orders:

splitting method	classical order
Lie splitting	1
Strang splitting	2
Yoshida splitting	4

It is clear that we can write down each splitting scheme with interchanged roles of A and B. This does not change the classical order and usually also not the convergence order of the special situation the splitting scheme is applied to. For long time computations it can be that one choice is preferable to the other one, namely if one ordering gives a gain in computing time by combining the last sub-step of one execution of the scheme with the first sub-step of the next execution. For instance, for the nonlinear Schrödinger equation it is advisable to choose A to be (a multiple of) the Laplace operator and B to be the nonlinearity when using the Strang splitting, see Section 2.1.

1.3.3. ADI splitting methods

The simplest numerical methods for solving differential equations are the explicit and the implicit Euler method. Let u_n^A and u_n^B be the numerical solutions after n time steps of length τ of the subproblems (1.3). Then the result after a further time step of length τ starting from them is

$$u_{n+1}^A = (I + \tau A)u_n^A \qquad \text{and} \qquad u_{n+1}^B = (I + \tau B)u_n^B,$$

respectively, for the explicit Euler method, and

$$u_{n+1}^A = (I - \tau A)^{-1}u_n^A \qquad \text{and} \qquad u_{n+1}^B = (I - \tau B)^{-1}u_n^B,$$

respectively, for the implicit Euler method. These methods are role models for all explicit and implicit methods since they show the typical properties of them.

A single step with an explicit method is very efficient but explicit methods have the disadvantage that they come along with a time step size restriction. The reason is that for partial differential operators A and B the explicit method is unstable for large time steps. The time step size restriction is of the type $\tau \leq cN_s^{-D}$—assuming a uniform space grid—, where N_s is the number of space discretization points in each direction and D the order of the differential equation. The necessity to make small time steps has the impact that many time steps have to be done, which causes a large total computation time. The complication is particularly severe for problems in higher dimensions since then the computation of one time step needs more time.

Implicit methods do in general not suffer from a time step size restriction but while applying them we have to solve a large system of equations, which usually needs a lot of computation time. This is especially a difficulty for multidimensional problems since the number of unknowns is proportional to N_s^d.

One remedy to these difficulties is to use so-called *alternating direction implicit* (*ADI*) splitting schemes, see [63]. We explain the idea using the example of the two-dimensional heat equation $\partial_t u = \partial_{xx} u + \partial_{yy} u$ (with suitable boundary conditions). We introduce two different numerical sub-methods and combine them to an ADI method. First, the second derivative in x-direction is computed implicitely and the second derivative in y-direction is computed explicitly. In the second sub-method it is done the other way around. These two methods are then executed after another (with the same time step size). In [63] it is shown that the resulting method is stable.

Transferring this idea to the abstract problem (1.2), we introduce the following splitting scheme. For a time step size $\tau > 0$ the result after the $(n+1)$-st time step is computed from the result after the n-th time step by

$$u_{n+1} = (I - \tfrac{\tau}{2}B)^{-1}(I + \tfrac{\tau}{2}A)\Big[(I - \tfrac{\tau}{2}A)^{-1}(I + \tfrac{\tau}{2}B)u_n\Big],$$

where we assume that $I - \frac{\tau}{2}A$ and $I - \frac{\tau}{2}B$ are invertible for all τ sufficiently small.

From now on we allow that the problem we investigate contains an inhomogeneity. This means that we look at a differential equation of the form

$$\begin{cases} u'(t) = Au(t) + Bu(t) + f(t), \quad t \geq t_0, \\ u(t_0) = u_0, \end{cases} \tag{1.5}$$

with a continuous function f. Inspired by the integrated form of (1.5),

$$u(t) = u_0 + \int_{t_0}^{t} (A + B)u(s)\,\mathrm{d}s + \int_{t_0}^{t} f(s)\,\mathrm{d}s,$$

we incorporate the impact of f into the numerical scheme by a quadrature rule. Choosing the trapezoidal rule, we define for a sufficiently small time step size $\tau > 0$ the splitting scheme

$$u_{n+1} = (I - \tfrac{\tau}{2}B)^{-1}(I + \tfrac{\tau}{2}A)\Big[(I - \tfrac{\tau}{2}A)^{-1}(I + \tfrac{\tau}{2}B)u_n + \tfrac{\tau}{2}\big(f(t_n) + f(t_{n+1})\big)\Big].$$

It is possible to choose other quadrature rules for the incorporation of the inhomogeneity. The most obvious alternative is the midpoint rule since it has the same order two and needs only one evaluation of f per time step. We did not work out the proofs in detail but we assume that this change does not affect the convergence orders we get in Chapter 9. At first sight the midpoint rule seems to be superior to the trapezoidal rule since it needs only one evaluation of the inhomogeneity per time step instead of two. But

the trapezoidal rule can compensate that by storing the evaluation of f for the next time step.

In principle it is also possible to use other quadrature rules than these two but this is not advisable. Using one of the two lower-order rectangular rules unfortunately reduces the overall convergence order of the scheme. The choice of a higher-order quadrature rule is a waste of computation time since gaining a higher convergence order than in our results would still be impossible due to the chosen arrangement of the operators A and B.

1.4. Tools from functional analysis and semigroup theory

In this section we state several important classical theorems from analysis that we use in this thesis.

1.4.1. Results from functional analysis

The first theorem gives a unique weak solution of linear partial differential equations, see Theorem 6.2.1 in [24].

Theorem 1.2 (Lemma of Lax–Milgram). *Let $(H, \|\cdot\|_H)$ be a real Hilbert space and $B : H \times H \to \mathbb{R}$ a bilinear mapping, for which there exist constants $\alpha, \beta > 0$ such that*

$$|B(u,v)| \leq \alpha \|u\|_H \|v\|_H$$

for all $u, v \in H$ and

$$B(u,u) \geq \beta \|u\|_H^2$$

for all $u \in H$. Furthermore, let $f : H \to \mathbb{R}$ be a bounded linear function. Then there exists a unique $u \in H$ such that

$$B(u,v) = f(v)$$

for all $v \in H$.

For some norm estimates it is crucial to have embeddings from some Sobolev spaces into Lebesgue spaces or spaces of continuous functions.

Theorem 1.3 (Sobolev embedding theorem). *Let Ω be a Lipschitz domain in $\mathbb{R}^d$ and $m \in \mathbb{N}$.*

(a) If $m > d/p$, then $W^{m,p}(\Omega) \hookrightarrow L^q(\Omega)$ for $q \in [p, \infty]$.

(b) If $m > d/p$ and Ω is a bounded cuboid, then $W^{m,p}(\Omega) \hookrightarrow C(\overline{\Omega})$.

(c) If $m = d/p$, then $W^{m,p}(\Omega) \hookrightarrow L^q(\Omega)$ for $q \in [p, \infty)$.

(d) If $m < d/p$, then $W^{m,p}(\Omega) \hookrightarrow L^q(\Omega)$ for $q \in [p, pd/(d - mp)]$.

(e) If $p = 2$, then the statements (a), (c) and (d) also holds for $\Omega = \mathbb{T}^d$.

PROOF:
For the case of a Lipschitz domain in $\mathbb{R}^d$ the statements and some more can be found in Theorem 4.12 in [1]. For the case of the torus and $p = 2$ part (a) follows from

$$\|f\|_\infty = \left\|\mathcal{F}^{-1}\mathcal{F}f\right\|_\infty \leq c\,\|\mathcal{F}f\|_{L^1} \leq c\,\left\|(1+|\cdot|^2)^{-s/2}\right\|_{L^2}\|f\|_{H^s} \leq c\,\|f\|_{H^s}\,, \tag{1.6}$$

and for part (d) see e.g. Corollary 1.2 in [8]. □

These Sobolev embeddings yield for up to three space dimensions in particular the following embeddings.

Corollary 1.4. *Let $d \in \{1, 2, 3\}$ and let either $\Omega \in \{\mathbb{R}^d, \mathbb{T}^d\}$ or $\Omega \subseteq \mathbb{R}^3$ a Lipschitz domain. Then $H^2(\Omega) \hookrightarrow L^\infty(\Omega)$ and $H^1(\Omega) \hookrightarrow L^6(\Omega)$.*

1.4.2. Results from semigroup theory

In this subsection we collect some theorems on strongly continuous semigroups, also called C_0-semigroups. An introduction into this topic and more detailed information can be found in [23].

The most important semigroups for this thesis are semigroups of contractions. The first theorem gives sufficient conditions under which a linear operator generates a C_0-semigroup of contractions, compare Theorem II.3.15 in [23].

Theorem 1.5 (Theorem of Lumer–Phillips). *Let X be a Banach space and let the operator $A : D(A) \subseteq X \to X$ be linear, closed, densely defined and dissipative. If the range of $\lambda I - A$ is dense in X for some $\lambda > 0$, then A generates a C_0-semigroup of contractions.*

Under suitable smallness assumptions, a perturbation of a generator of a semigroup of contractions is again a generator. Theorem III.2.7 in [23] provides the following result on perturbation by a bounded and dissipative operator.

Theorem 1.6 (Theorem of dissipative perturbation). *Let X be a Banach space, $A : D(A) \subseteq X \to X$ generate a C_0-semigroup of contractions and $B \in \mathcal{B}(X)$ be dissipative. Then $A + B$ generates a C_0-semigroup of contractions on $D(A)$.*

The following theorem yields a further statement on the generation of a semigroup of contractions, see Theorem II.3.5 in [23]. It can also be used the other way around to get from a semigroup of contractions a resolvent estimate.

Theorem 1.7 (Theorem of Hille–Yosida). *Let X be a Banach space and $A : D(A) \subseteq X \to X$ a linear operator. Then the following properties are equivalent.*

(a) A generates a C_0-semigroup of contractions.

(b) A is closed, densely defined, every $\lambda > 0$ belongs to the resolvent set of A, and one has the estimate $\|(\lambda I - A)^{-1}\|_{\mathcal{B}(X)} \leq \frac{1}{\lambda}$ for all $\lambda > 0$.

On a Hilbert space one can characterize the generators of unitary groups by the following result, see Theorem II.3.24 in [23].

Theorem 1.8 (Stone's Theorem). *Let H be a Hilbert space and $A : D(A) \subseteq H \to H$ a densely defined operator. Then A generates a C_0-group of unitary operators if and only if A is skew-adjoint.*

The main purpose of semigroups for this thesis is that they are closely connected to the solutions of Cauchy problems.

Theorem 1.9. *Let X be a Banach space, $A : D(A) \subseteq X \to X$ a linear operator that generates a C_0-semigroup $(T(t))_{t\geq 0}$, $u_0 \in D(A)$ and $f \in C^1([0,\infty), X) + C([0,\infty), D(A))$.*

(a) The inhomogeneous Cauchy problem

$$\begin{cases} u'(t) = Au(t) + f(t), & t \geq 0, \\ u(0) = u_0, \end{cases}$$

has a unique solution u in $C^1([0,\infty), X) \cap C([0,\infty), D(A))$, which satisfies

$$u(t) = T(t)u_0 + \int_0^t T(t-s)f(s)\,\mathrm{d}x = T(t)u_0 - \int_0^t T(s)f(t-s)\,\mathrm{d}s \tag{1.7}$$

for $t \geq 0$. Moreover,

$$\|u\|_{C([0,T],D(A))\cap C^1([0,T],X)} \leq c\big(\|u_0\|_{D(A)} + \|f\|_{C^1([0,T],X)+C([0,T],D(A))}\big)$$

for all $T > 0$.

(b) Let $T > 0$. If $u_0 \in D(A^2)$ and $f \in C^2([0,T], X) \cap C([0,T], D(A))$, then u belongs to $C([0,T], D(A^2))$ and we have

$$\|u\|_{C([0,T],D(A^2))} \leq c\big(\|u_0\|_{D(A^2)} + \|f\|_{C([0,T],D(A))} + \|f\|_{C^2([0,T],X)}\big).$$

PROOF:
For the statements of part (a) compare Section VI.7a in [23], and see Corollary 4.2.5 and 4.2.6 in [62] for the solution formula. The proofs of the two corollaries further shows the estimate of (a).

For the proof of part (b) let $T > 0$. We differentiate (1.7) and get

$$u'(t) = T(t)\big(Au_0 + f(0)\big) + \int_0^t T(s) f'(t-s)\,\mathrm{d}s$$

for all $t \geq 0$. Since $Au_0 + f(0) \in D(A)$ and $f \in C^1([0,T],X)$, Corollary 4.2.5 and 4.2.6 in [62] then yield $u' \in C^1([0,T],X) \cap C([0,T],D(A))$. From $Au'(t) = A^2u(t) + Af(t)$ for all $t \geq 0$ we infer that $A^2u = Au' - Af$ belongs to $C([0,T],X)$. This gives with part (a) and the proofs of Corollary 4.2.5 and 4.2.6 in [62] that

$$\begin{aligned}
&\|u\|_{C([0,T],D(A^2))} \\
&\quad = \|u\|_{C([0,T],X)} + \big\|A^2u\big\|_{C([0,T],X)} \\
&\quad \leq \|u\|_{C([0,T],X)} + \|Au'\|_{C([0,T],X)} + \|Af\|_{C([0,T],X)} \\
&\quad \leq c\big(\|u_0\|_{D(A)} + \|f\|_{C^1([0,T],X)} + \|u_0\|_{D(A^2)} + \|f\|_{C^2([0,T],X)} + \|f\|_{C([0,T],D(A))}\big),
\end{aligned}$$

which is the desired estimate. □

Finally, we recall Sobolev spaces of negative orders associated to a semigroup. The statements can be found in Section II.5a in [23] and in Theorem V.1.4.6 in [2].

Proposition 1.10. *Let X be a Banach space and $A : D(A) \subseteq X \to X$ a linear operator that generates a C_0-semigroup $(e^{tA})_{t\geq 0}$ on X. Then there exists a $\lambda > 0$ in the resolvent set of A. We define the Sobolev space of order -1 associated to the semigroup $(e^{tA})_{t\geq 0}$,*

$$X_{-1}^A := \Big(X, \big\|(\lambda I - A)^{-1}\cdot\big\|_X\Big)^{\sim}, \tag{1.8}$$

which denotes the completion of X with respect to the norm $\big\|(\lambda I - A)^{-1}\cdot\big\|_X$. The operator A has an extension $A_{-1} : X \to X_{-1}^A$ that generates an extended C_0-semigroup on X_{-1}^A. Inductively, the Sobolev spaces of order -2 and so on are defined. Furthermore,

$$X_{-1}^{A^*} \cong D\big(A^{**}\big)^* = D(A)^* \qquad \text{and} \qquad X_{-2}^{A^*} \cong D(A^2)^*.$$

Part II.

The Strang and the Lie splitting for the cubic nonlinear Schrödinger equation

2. Basic properties of the nonlinear Schrödinger equation

In this chapter we introduce the cubic nonlinear Schrödinger equation (NLS) and provide the background for this part of the thesis. In Section 2.1 we discuss the problem and our splitting schemes, followed by a summary of the state of the art and an outline of our theorems. In Section 2.2 we describe the needed functional analytic setting and prove auxiliary lemmas on function spaces, the free Schrödinger group and the solutions to the cubic NLS.

2.1. The nonlinear Schrödinger equation and the splitting schemes

Among the many different *semilinear Schrödinger equations* we focus on the one with a cubic nonlinearity. Let $\mu \in \{-1, 1\}$ be a parameter and $d \in \{1, 2, 3\}$ be the spatial dimension. Let the spatial domain Ω be either the full space $\mathbb{R}^d$ or the d-dimensional torus $\mathbb{T}^d$. We choose the initial time $t_0 = 0$ and restrict ourselves to non-negative times. The *cubic nonlinear Schrödinger equation* then reads

$$\begin{cases} \partial_t u(t) = i\Delta u(t) - i\mu \left|u(t)\right|^2 u(t), \quad t \geq 0, \\ u(0) = u_0, \end{cases} \tag{2.1}$$

for a given initial function $u_0 \in H^2(\Omega)$. The parameter μ determines the sign of the nonlinearity. In the *focusing case* $\mu = -1$ the problem (2.1) has blow-up solutions for $d \geq 2$, see e.g. Theorem 6.5.10 in [13]. In contrast to this the solutions obtained in the *defocusing case* $\mu = 1$ are global in time by e.g. Corollary 6.1.2 in [13] for $\Omega = \mathbb{R}^d$ and Section V.2 in [11] for $\Omega = \mathbb{T}^d$. From now on we omit Ω in our notation if we do not need to distinguish between $\mathbb{R}^d$ and $\mathbb{T}^d$.

The cubic nonlinear Schrödinger equation arises in nonlinear optics and in the theory of shallow water waves as amplitude equation that approximatly determines the evolution of wave packets. A variant of (2.1) with a potential term is the Gross–Pitaevskii equation that governs the behaviour of Bose–Einstein condensates. Further information on the

physical background can be found in [55] and [67]. Semilinear Schrödinger equations are investigated in the monograph [13] in great detail and generality.

We consider (2.1) as an equation in L^2 and thus require that the initial function is at least in H^2. Due to Theorem 4.1 and 4.2 in [46] we have for all $u_0 \in H^s$ with $s \geq 2$ a unique local *H^s-solution*. This is a function

$$u \in C^1([0, T_{\max}), H^{s-2}) \cap C([0, T_{\max}), H^s)$$

fulfilling (2.1), where $T_{\max} \in (0, \infty]$ is the *maximal existence time*. The blow-up alternative says that the solution exists either for all times, which means $T_{\max} = \infty$, or that the norm of the solution tends to infinity at a finite time that we then call $T_{\max} < \infty$. Because we want to guarantee the existence of the solution up to the end time of our observation interval, we restrict the solution to a time interval $[0, T]$ with a fixed finite *end time* $T < T_{\max}$. As mentioned above we can choose T arbitrarily large if $\mu = 1$. This is also the case if $d = 1$, see Section 6.6 in [13]. All these properties on the maximal existence times hold also true for negative times, which we do not consider in this thesis.

There are two reasons why we restrict ourselves to at most three space dimensions. First, in the physically most relevant situations we have one, two or three dimensions. Second, replacing H^2 by H^k with $k > 2$ we could treat great higher dimensions than three since we then have the needed Sobolev embeddings but we omit this for the simplicity of the presentation. In the case of only one or two spatial dimensions some simplifications of the following proofs are possible, which we do not discuss.

Apart from the cubic nonlinearity it is also of interest to look at the more general equation

$$\begin{cases} \partial_t u(t) = i\Delta u(t) - i\mu\varphi\big(|u(t)|^2\big)u(t), \quad t \geq 0, \\ u(0) = u_0 \in H^2, \end{cases}$$

with a twice continuously differentiable function $\varphi : \mathbb{R} \to \mathbb{R}$ with $\varphi(0) = 0$. Nevertheless, the cubic nonlinearity is its most important representative since it appears in many applications and since it can be considered as a model case for the more general situations. The analysis in this thesis is flexible enough to be extended to the general nonlinearities φ. However, to avoid technicalities in the context of fractional Sobolev spaces, we restrict ourselves to the cubic case.

The solution to the nonlinear ordinary differential equation

$$\partial_t u(t) = -i\mu\,|u(t)|^2\,u(t)$$

with initial value $u(0) = u_0$ is given for all $t \geq 0$ by the simple formula

$$u(t) = \exp\big(-i\mu t\,|u_0|^2\big)u_0.$$

The linear equation

$$\partial_t u(t) = i\Delta u(t)$$

can easily be solved by means of the Fourier transform, which can numerically be approximated efficiently on the torus. These observations are exploited when using the Lie and the Strang splitting scheme for (2.1). In the Lie splitting scheme the numerical solution after one time step $\tau > 0$ starting at $u_0 \in H^2$ is given by

$$\Phi_\tau(u_0) := \exp\bigl(-i\mu\tau\,|\widetilde{u}|^2\bigr)\widetilde{u} \qquad \text{with} \qquad \widetilde{u} := T(\tau)u_0, \tag{2.2}$$

and in the Strang splitting scheme by

$$\begin{aligned} \Psi_\tau(u_0) &:= T(\tau/2)u^{**} \\ \text{with} \qquad u^{**} &:= \exp\bigl(-i\mu\tau\,|u^*|^2\bigr)u^* \qquad \text{with} \qquad u^* := T(\tau/2)u_0, \end{aligned} \tag{2.3}$$

where $T(\cdot)$ denotes the *free Schrödinger group*.

We could interchange the usage of the solution formula for the linear and the nonlinear equation in these splitting schemes. In applications one is sometimes only interested in the value of the solution at the end time. For the Strang splitting this means that the last sup-step of each execution of the scheme and the first sub-step of the next execution can be combined in the computation. Calculating the fast Fourier transform and its inverse in the computation of the solution of the linear equation takes much more computing time than evaluating the action of a multiplication operator in the solution formula for the nonlinear equation. Therefore, it is not advisable for the Strang splitting to interchange the usage of the solution formulas.

C. Lubich showed in [51] second-order convergence in time of the Strang splitting scheme for initial functions in $H^4(\mathbb{R}^d)$ with a proof based on the theory of Lie derivatives (see also [43] for linear Schrödinger equations). More precisely, there exists a bound $\tau_0 \in (0, T]$ on the time step size such that

$$\|u(n\tau) - \Psi_\tau^n(u_0)\|_{L^2} \leq C\tau^2$$

for all $u_0 \in H^4(\mathbb{R}^d)$, $\tau \in (0, \tau_0]$ and $n \in \mathbb{N}$ with $n\tau \in [0, T]$ with a constant $C \geq 0$ depending only on T and on the norm of u in $C([0, T], H^4(\mathbb{R}^d))$, see Theorem 7.1 in [51] and also [40]. We give a complete proof of this theorem in Chapter 3. We note that in [51] the time step size restriction was missing. In the later paper [40], coauthored by C. Lubich, this was then elaborated in a somewhat different context.

For smooth solutions a Taylor series expansion shows that the Lie and the Strang splitting are of classical order one and two, respectively, cf. Section 1.3.2. Hence, more regular initial functions do not lead to a higher convergence order. Higher-order splitting methods for Schrödinger equations were investigated in [68] and in [6], and in [69] for the Gross–Pitaevskii equation.

In our Theorem 4.1 we reduce the level of regularity of the initial functions and therefore of the solutions to $H^{2+2\theta}$ with $\theta \in (0,1)$. We show an error estimate in L^2 of the Strang splitting with the corresponding fractional convergence order $1+\theta$. Afterwards we prove an analogous fractional error estimate which shows the first order convergence in L^2 for the Strang and the Lie splitting for initial functions in H^2, see Theorem 5.1 and 5.2. These three theorems have been published in [22]. Results for the Lie splitting in the case of the cubic NLS have been known so far only in spaces of functions on the torus with summable Fourier coefficients. For this see Proposition IV.6 of [25], where the calculus of Lie derivatives was used. Moreover, for nonlinearities of the type $i\lambda|u|^p u$ with $\lambda \in \mathbb{R}$ and $p < 4/3$ first-order convergence of the Lie splitting in L^2 was shown in [42] for initial functions in H^2 by different methods than ours. In the thesis at hand we focus on the time integration and do not treat the space discretization (which was studied in e.g. [25]).

The strategy for the proof of the theorems is to show a local error bound and that the numerical solution after one time step $\tau > 0$ is Lipschitz continuous with respect to the initial function. To iterate this stability estimate, the Lipschitz constant has to be of the form $e^{c\tau}$. One then obtains a Lipschitz bound on time intervals $[0, n\tau]$ with constant $e^{cn\tau}$ for $n \in \mathbb{N}$. Because of the nonlinearity, c depends on the (so far uncontrolled) H^s-norm of the numerical solution on $[0, n\tau]$, cf. Lemma 3.4, 5.5 and 5.6. Here we take $s = 2$ for the case of initial functions in $H^{2+2\theta}$ and $s = 7/4$ for the case of initial functions in H^2. By means of a telescopic sum, see e.g. [34] or [40], we then deduce a global error bound in our Theorems 3.1, 4.1, 5.1 and 5.2. We measure the error in L^2, but we can also bound it in H^s (with a smaller fractional convergence order). Since the exact solution is bounded in H^s, the needed a priori estimate on the numerical solution in H^s follows under a time step size restriction, see [40] or our Lemmas 3.6, 4.3 and 5.7.

2.2. The functional analytic setting

We define the operators

$$A : H^2 \to L^2; \quad Au := i\Delta u, \qquad \text{and} \qquad B : H^2 \to L^2; \quad B(u) := -i\mu\,|u|^2\,. \tag{2.4}$$

They are the *splitting operators* we are going to use. The free Schrödinger group generated by A is denoted by $T(\cdot)$. The mapping $I - \Delta : H^{s+2} \to H^s$ is for all $s \geq 0$ an isomorphism. This fact can be deduced from the characterization (1.1) of the Sobolev spaces via the Fourier transform, using that the Laplace operator corresponds to the symbol $|\cdot|^2$ in Fourier space. One furthermore sees that Δ is self-adjoint in H^s for all $s \geq 0$. Hence, $i\Delta$ is skew-adjoint in H^s, so that the restriction of $T(\cdot)$ to H^s is a unitary C_0-group on H^s by Stone's Theorem for all $s \geq 0$. We denote these restrictions also by $T(\cdot)$. With the

introduced notation problem (2.1) takes the form

$$\begin{cases} \partial_t u(t) = Au(t) + B(u(t))u(t), \quad t \geq 0, \\ u(0) = u_0 \in H^2. \end{cases} \tag{2.5}$$

We look at the two "subproblems"

$$\begin{cases} \partial_t v(t) = Av(t) = i\Delta v(t), \quad t \geq 0, \\ v(0) = v_0 \in H^2, \end{cases} \tag{2.6}$$

and

$$\begin{cases} \partial_t w(t) = B(w(t))w(t) = -i\mu \left|w(t)\right|^2 w(t), \quad t \geq 0, \\ w(0) = w_0 \in H^2. \end{cases} \tag{2.7}$$

The subproblem (2.6) is uniquely solved by $v(t) = T(t)v_0$ and the subproblem (2.7) by

$$w(t) = e^{tB(w_0)}w_0, \tag{2.8}$$

both for all $t \geq 0$. For both subproblems we thus have explicitly given solution formulas. A fully discrete numerical approximation to the solution of subproblem (2.6) can efficiently be computed at least on the torus by using the fast Fourier transform, see e.g. [25]. The solution of subproblem (2.7) can quickly be calculated by means of the solution formula. Therefore, splitting methods like (2.2) and (2.3) are very attractive for the numerical treatment of (2.5). With the above notations, the Lie splitting (2.2) reads

$$\Phi_\tau(u_0) := \exp\big(\tau B(\widetilde{u})\big)\widetilde{u} \qquad \text{with} \qquad \widetilde{u} := T(\tau)u_0 \tag{2.9}$$

and the Strang splitting (2.3) becomes

$$\begin{aligned} \Psi_\tau(u_0) &:= T(\tau/2)u^{**} \\ \text{with} \qquad u^{**} &:= \exp\big(\tau B(u^*)\big)u^* \qquad \text{and} \qquad u^* := T(\tau/2)u_0. \end{aligned} \tag{2.10}$$

We recall the well-known fact that the space H^s is an *algebra* if $s > d/2$ and several related properties, which are crucial for our analysis.

Lemma 2.1. *(a) Let $s \in (3/2, \infty)$. Then the product of functions $f, g \in H^s$ belongs to H^s and satisfies*

$$\|fg\|_{H^s} \leq c \|f\|_{H^s} \|g\|_{H^s}.$$

The product of a function $f \in H^s$ and a function $g \in L^2$ belongs to L^2 and satisfies

$$\|fg\|_{L^2} \leq c \|f\|_{H^s} \|g\|_{L^2}.$$

(b) Let $s \in (3/2, \infty)$, $t \geq 0$, and $v, w \in H^s$ with $\|v\|_{H^s} \leq r$ and $\|w\|_{H^s} \leq r$. Then we have

$$\begin{aligned} \|B(v)\|_{H^s} &\leq c\,\|v\|_{H^s}^2 \leq cr^2, \\ \|B(v) - B(w)\|_{H^s} &\leq c\big(\|v\|_{H^s} + \|w\|_{H^s}\big)\,\|v - w\|_{H^s} \leq cr\,\|v - w\|_{H^s}\,. \end{aligned}$$

If $s \in [s_1, s_2] \subseteq (3/2, \infty)$, then all the constants only depend on s_1 and s_2.

Proof:
(a): Let $s > 3/2$ and $f, g \in H^s$. We have

$$\begin{aligned} \big(1 + |\xi|^2\big)^{s/2} &\leq c\Big(\big(1 + |\xi - \eta|^2\big) + \big(1 + |\eta|^2\big)\Big)^{s/2} \\ &\leq c\Big(\big(1 + |\xi - \eta|^2\big)^{s/2} + \big(1 + |\eta|^2\big)^{s/2}\Big) \end{aligned}$$

for all $\xi, \eta \in \mathbb{R}^d$, using basic estimates for the roots for $s \in (3/2, 2]$ and Hölder's inequality for $s > 2$. From this estimate and $\mathcal{F}(fg) = c(\mathcal{F}f) * (\mathcal{F}g)$ we derive that

$$\begin{aligned} \big(1 + |\xi|^2\big)^{s/2}\,|\mathcal{F}(fg)(\xi)| &\leq c \int_{\mathbb{R}^d} \big(1 + |\xi|^2\big)^{s/2}\,|(\mathcal{F}f)(\xi - \eta)(\mathcal{F}g)(\eta)|\,\mathrm{d}\eta \\ &\leq c\Big(\Big|\big(1 + |\cdot|^2\big)^{s/2}\mathcal{F}f\Big| * |\mathcal{F}g|\Big)(\xi) + c\Big(|\mathcal{F}f| * \Big|\big(1 + |\cdot|^2\big)^{s/2}\mathcal{F}g\Big|\Big)(\xi). \end{aligned}$$

Young's inequality for convolutions and the Sobolev embedding $H^s \hookrightarrow L^\infty$ in (1.6) thus yield

$$\begin{aligned} \|fg\|_{H^s} &\leq c\big(\|f\|_{H^s}\,\|\mathcal{F}g\|_{L^1} + \|\mathcal{F}f\|_{L^1}\,\|g\|_{H^s}\big) \\ &\leq c\big(\|f\|_{H^s}\,\|g\|_{L^\infty} + \|f\|_{L^\infty}\,\|g\|_{H^s}\big) \leq c\,\|f\|_{H^s}\,\|g\|_{H^s}\,. \end{aligned}$$

The rest of the statement follows directly from the Sobolev embedding. The statements of part (b) follow directly from part (a).

The constants are uniformly bounded for $s \in [s_1, s_2]$ since the Sobolev embedding constants satisfy this property and since the constants depend on s only via the Sobolev embeddings. □

Remark 2.2. *In the rest of our analysis we only deal with the case $s \in [7/4, 4]$, so that the constant c in the previous lemma can be chosen independently of s.*

Theorem 4.1 in [46] shows that for $u_0 \in H^s$ with $s \geq 2$ the problem (2.5) is locally wellposed, i.e. there exists a time $T > 0$ such that there exists a unique solution $u = u(\cdot, u_0) \in C([0, T], H^s)$ to (2.5). Throughout the thesis T is chosen in this way. (In the defocusing case $\mu = 1$ one obtains a global solution on $\mathbb{R}_+$, see e.g. Corollary 6.1.2 in [13]

for $\Omega = \mathbb{R}^d$ and Section V.2 in [11] for $\Omega = \mathbb{T}^d$, but we will not need this fact.) The solution is given by

$$\begin{aligned} u(t) &= T(t)u_0 - i\mu \int_0^t T(t-r)\big(|u(r)|^2 u(r)\big)\,\mathrm{d}r \\ &= T(t)u_0 + \int_0^t T(t-s)B(u(s))u(s)\,\mathrm{d}s, \end{aligned} \tag{2.11}$$

see Theorem refthm:solinhomCauchypbDuhamel. Because H^s is an algebra by Lemma 2.1, the function $|u|^2 u$ belongs to $C([0,T], H^s)$. Hence, by standard semigroup theory, u is contained in $C^1([0,T], H^{s-2})$ and solves problem (2.5) in H^{s-2}. Below we use the quantities

$$M_s := \sup_{t\in[0,T]} \|u(t)\|_{H^s} \qquad \text{for } s \geq 0,$$

whenever these expressions are finite. We remark that M_s depends only on u_0, s and T. By the representation of the Sobolov spaces via the Fourier transform, the Fourier transform is up to a constant an isometric isomorphism from H^s to the weighted Lebesgue space

$$L^2_s := \big\{ f \in L^2 \;\big|\; \big(1+|x|^2\big)^{s/2} |f(x)| \in L^2 \big\}.$$

For all $0 < s_1 < s_2$ we infer from $\big(1+|x|^2\big)^{s_1/2} \leq \big(1+|x|^2\big)^{s_2/2}$ that

$$\|f\|_{H^{s_1}} = \|\mathcal{F}f\|_{L^2_{s_1}} \leq \|\mathcal{F}f\|_{L^2_{s_2}} = \|f\|_{H^{s_2}}$$

for all $f \in H^{s_2}$. The equivalence of the two Sobolev norms in Section 1.1 thus implies that $M_{s_1} \leq cM_{s_2}$ for all $0 \leq s_1 \leq s_2 \leq 4$ for a constant c independent of s_1 and s_2.

We close this section by stating several important regularity properties of the free Schrödinger group and the solutions to (2.5).

Lemma 2.3. *Let $\eta \in (0,1)$ and $s \geq 0$.*

(a) For $f \in H^{2\eta}$ and $g \in H^2$, we have $fg \in H^{2\eta}$ and

$$\|fg\|_{H^{2\eta}} \leq c\,\|f\|_{H^{2\eta}}\,\|g\|_{H^2}\,.$$

(b) For each $y \in H^{s+2\eta}$, the mapping $T(\cdot)y : [0,\infty) \to H^s$ is η-Hölder continuous with

$$\|T(t_1)y - T(t_2)y\|_{H^s} \leq c\,|t_1 - t_2|^\eta\,\|y\|_{H^{s+2\eta}}$$

for all $t_1, t_2 \geq 0$.

(c) Let $s > 3/2$. For each $y \in H^{s+2\eta}$, the solution $u(\cdot, y) : [0,T] \to H^{s+2\eta}$ to (2.5) is η-Hölder continuous on H^s with

$$\begin{aligned} \|u(t_1,y) - u(t_2,y)\|_{H^s} &\leq c\big(M_{s+2\eta} + M_s^3 T^{1-\eta} + T M_{s+2\eta}^3\big)\,|t_1 - t_2|^\eta \\ &=: C(M_{s+2\eta}, T)\,|t_1 - t_2|^\eta \end{aligned}$$

for all $t_1, t_2 \in [0,T]$.

The above constants c do not depend on η.

PROOF:
Let $\eta \in (0,1)$. We first recall that $H^{s+2\eta}$ is an interpolation space between H^s and H^{s+2} by Theorem 5.4.1 in [9] in combination with the Fourier transform. (See also Theorem 6.2.4 and 6.4.4 in [9] for $\mathbb{R}^d$.) We observe that the constants involved in this proof can be chosen independently of η.

(a) Let $g \in H^2$. The norms of the linear operators $V_1 : L^2 \to L^2$ and $V_2 : H^2 \to H^2$ given by $V_j f := fg$ for $j = 1, 2$ are bounded by $c \|g\|_{H^2}$ due to Lemma 2.1. Assertion (a) then follows by interpolation.

(b) Let $t_1, t_2 \geq 0$ with $t_1 < t_2$ be fixed. We look at the linear mapping $\widetilde{T}_{t_1,t_2} : H^s \to H^s$; $\widetilde{T}_{t_1,t_2} y := T(t_1)y - T(t_2)y$, whose norm is bounded by 2. We also use its restriction $\widetilde{T}_{t_1,t_2} : H^{s+2} \to H^s$. For $y \in H^{s+2}$, we have $\frac{\mathrm{d}}{\mathrm{d}t} T(t)y = T(t)Ay$ and hence

$$\left\| \widetilde{T}_{t_1,t_2} y \right\|_{H^s} \leq \sup_{t \in [t_1,t_2]} \|T(t)Ay\|_{H^s} \, |t_1 - t_2| \leq c \, |t_1 - t_2| \, \|y\|_{H^{s+2}} \, .$$

Interpolation now yields assertion (b).

(c) The representation (2.11), part (b), the unitarity of $T(\cdot)$ on H^s and Lemma 2.1 imply

$$\begin{aligned}
&\|u(t_1, y) - u(t_2, y)\|_{H^s} \\
&\quad \leq \|T(t_1)y - T(t_2)y\|_{H^s} + \int_{t_1}^{t_2} \left\| T(t_2 - s)\big(|u(s)|^2 \, u(s)\big) \right\|_{H^s} \mathrm{d}s \\
&\qquad + \int_0^{t_1} \left\| \big(T(t_2 - t_1) - I\big) T(t_1 - s)\big(|u(s)|^2 \, u(s)\big) \right\|_{H^s} \mathrm{d}s \\
&\quad \leq c \, |t_1 - t_2|^\eta \, \|y\|_{H^{s+2\eta}} + c M_s^3 \, |t_1 - t_2|^\eta \, T^{1-\eta} + c \, |t_1 - t_2|^\eta \, T M_{s+2\eta}^3
\end{aligned}$$

for all $0 \leq t_1 \leq t_2 \leq T$. □

3. Convergence of the Strang splitting for initial functions in H^4

In [51], C. Lubich showed that the Strang splitting applied to the cubic NLS on $\mathbb{R}^d$ converges with order two if the initial function is contained in H^4. We present the main theorem and some auxiliary properties of the splitting scheme in Section 3.1. The proof contains estimates in H^2 and L^2, which are presented in the Sections 3.2 and 3.3, respectively. The ideas for the intermediate results and the structure of the proof are taken from [51] and [40]. We give the full prove of this theorem since the original one is a bit sketchy and since it provides the background for our later results.

3.1. The convergence theorem for initial functions in H^4

The following convergence theorem for the Strang splitting can be found for the full-space situation as Theorem 7.1 in [51].

Theorem 3.1. *For each $u_0 \in H^4$ there exists a bound $\tau_0 > 0$ on the time step size such that we have*

$$\|u(n\tau) - \Psi_\tau^n(u_0)\|_{L^2} \leq C\tau^2$$

for all $\tau \in (0, \tau_0]$ and $n \in \mathbb{N}$ with $n\tau \leq T$ with a constant $C \geq 0$ that depends only on u_0 and T. More precisely, C depends only on T and M_4.

The number $\tau_0 = \tau_0(T, M_2)$ is given in Lemma 3.6.

Remark 3.2. *The dependency of C on M_4 in the previous theorem shows that the error constant is large for rapidly oscillating solutions and therefore also for rapidly oscillating initial functions. We confirm this numerically in Section 6.5.*

We first show that the local error in H^2 is of order two.

Lemma 3.3. *For all $u_0 \in H^4$ and $\tau \in (0, T]$ we have*

$$\|u(\tau) - \Psi_\tau(u_0)\|_{H^2} \leq C_1\tau^2,$$

with a constant $C_1 \geq 0$ depending only on M_4.

3. Convergence of the Strang splitting for initial functions in H^4

We next need a *stability lemma* for the Strang splitting.

Lemma 3.4. *Let $M \geq 0$ and $u_0, v_0 \in H^2$ with $\|u_0\|_{H^2} \leq M$ and $\|v_0\|_{H^2} \leq M$. Then there exists a constant $C_2 \geq 0$, only depending on M, such that*

$$\|\Psi_\tau(u_0) - \Psi_\tau(v_0)\|_{H^2} \leq e^{C_2\tau} \|u_0 - v_0\|_{H^2}$$

for all $\tau \in (0, T]$.

Here, the precise form of the constant in the estimate is crucial since its n-th power enters in the proof of Theorem 3.1. In this proof we also need the following property of the numerical scheme.

Definition 3.5. *Let $T > 0$, $s \geq 2$ and ϕ_τ be a time integration scheme. We call the scheme ϕ_τ* strongly bounded for (2.5) in H^s for initial functions in H^t *with time step size bound $\tau_0 \in (0, T]$ if for all initial functions $u_0 \in H^t$ there exists a constant $\widehat{C} \geq 0$, only depending on u_0 and T, such that for all $\tau \in (0, \tau_0]$, $n \in \mathbb{N}$ with $n\tau \leq T$ and $k \in \{0, \dots, n\}$ we have $\left\|\phi_\tau^{n-k}(u(k\tau))\right\|_{H^s} \leq \widehat{C}$. Here, u denotes the solution to (2.5) with initial function u_0.*

The Strang splitting for the cubic NLS is strongly bounded in H^2.

Lemma 3.6. *Let $u_0 \in H^4$. Then there exists a bound $\tau_0 > 0$ on the time step size, which is given by*

$$\tau_0 := \min\left\{\frac{M_2}{Te^{TC_2}C_1}, T\right\},$$

with C_1 from Lemma 3.3 and C_2 from Lemma 3.4, such that the following two statements hold true.

(a) For all $\tau \in (0, \tau_0]$ and $n \in \mathbb{N}$ with $n\tau \leq T$, we have

$$\|\Psi_\tau^n(u_0) - u(n\tau)\|_{H^2} \leq C\tau,$$

with a constant $C \geq 0$ depending only on T and M_2, i.e. the Strang splitting converges in H^2 with order one.

(b) Ψ_τ is strongly bounded for (2.5) in H^2 for initial functions in H^4, i.e. there exists a constant $\widehat{C} \geq 0$, only depending on T and M_2, such that $\left\|\Psi_\tau^{n-k}(u(k\tau))\right\|_{H^2} \leq \widehat{C}$ for all $\tau \in (0, \tau_0]$ and $n \in \mathbb{N}$ with $n\tau \leq T$ and $k \in \{0, \dots, n\}$. In particular, the numerical solution is bounded in H^2 (choose $k = 0$).

The above lemmas are proved in Section 3.2. In the next lemma we show that the local error in L^2 is of order three, i.e. one order higher than in H^2.

Lemma 3.7. *For all $u_0 \in H^4$ and $\tau \in (0,T]$ we have*

$$\|u(\tau) - \Psi_\tau(u_0)\|_{L^2} \leq C_3 \tau^3,$$

with a constant $C_3 \geq 0$ depending only on M_4.

Due to the nonlinearity, we obtain a weaker stability property in L^2 than in H^2, which we call H^2-*conditional stability*. For this reason we have to invoke the strong boundedness in H^2. It is used to apply Lady Windermere's fan, see [34], in the proof of Theorem 3.1.

Lemma 3.8. *Let $M \geq 0$ and $u_0, v_0 \in H^2$ with $\|u_0\|_{H^2} \leq M$ and $\|v_0\|_{H^2} \leq M$. Then there exists a constant $C_4 \geq 0$, only depending on M, such that*

$$\|\Psi_\tau(u_0) - \Psi_\tau(v_0)\|_{L^2} \leq e^{C_4 \tau} \|u_0 - v_0\|_{L^2}$$

for all $\tau \in (0,T]$.

The preceding two lemmas and Theorem 3.1 are shown in Section 3.3.

3.2. The estimate in H^2

We prove Lemma 3.3 and 3.4 and combine them to show Lemma 3.6.

3.2.1. The local error in the H^2-norm

Proof (of Lemma 3.3):
Let $u_0 \in H^4$ and $\tau > 0$. By (2.11), the solution to problem (2.5) at time τ is given by

$$u(\tau) = T(\tau)u_0 + \int_0^\tau T(\tau - s)B(u(s))u(s)\,\mathrm{d}s.$$

Plugging this formula into itself, we derive the representation

$$\begin{aligned} u(\tau) = T(\tau)u_0 &+ \int_0^\tau T(\tau - s)B(u(s))T(s)u_0\,\mathrm{d}s \\ &+ \int_0^\tau T(\tau - s)B(u(s)) \int_0^s T(s - \sigma)B(u(\sigma))u(\sigma)\,\mathrm{d}\sigma\,\mathrm{d}s, \end{aligned} \tag{3.1}$$

which is valid in H^2. To show a corresponding formula for the numerical approximation, we use the Taylor expansion

$$e^{\tau x} = 1 + \tau x + \int_0^\tau (\tau - s)x^2 e^{sx}\,\mathrm{d}s.$$

Applying this to $u^{**} = e^{\tau B(u^*)}u^*$, we infer

$$u^{**} = u^* + \tau B(u^*)u^* + \int_0^\tau (\tau - s)B(u^*)^2 e^{sB(u^*)}u^*\,\mathrm{d}s.$$

Because $\Psi_\tau(u_0) = T(\tau/2)u^{**}$ and $u^* = T(\tau/2)u_0$, see (2.10), the numerical solution after one time step is thus given by

$$\Psi_\tau(u_0) = T(\tau)u_0 + \tau T(\tau/2)B(u^*)T(\tau/2)u_0 + \int_0^\tau (\tau - s)T(\tau/2)B(u^*)^2 e^{sB(u^*)}T(\tau/2)u_0 \,\mathrm{d}s.$$

This equation and the representation (3.1) yield

$$\begin{aligned} u(\tau) - \Psi_\tau(u_0) &= \left(\int_0^\tau T(\tau - s)B(u(s))T(s)u_0 \,\mathrm{d}s - \tau T(\tau/2)B(u^*)T(\tau/2)u_0 \right) \\ &\quad + \left(\int_0^\tau T(\tau - s)B(u(s)) \int_0^s T(s - \sigma)B(u(\sigma))u(\sigma) \,\mathrm{d}\sigma \,\mathrm{d}s \right. \\ &\quad \left. - \int_0^\tau (\tau - s)T(\tau/2)B(u^*)^2 e^{sB(u^*)}T(\tau/2)u_0 \,\mathrm{d}s \right) \qquad (3.2) \\ &=: I_1 + I_2. \end{aligned}$$

1) *Bound on* I_1: We look at the function $w : [0, T] \to H^2$ defined by

$$w(s) := T(\tau - s)B(u(s))T(s)u_0.$$

We then estimate with the midpoint quadrature rule

$$\begin{aligned} \|I_1\|_{H^2} &\le \left\| \int_0^\tau w(s) \,\mathrm{d}s - \tau w(\tau/2) \right\|_{H^2} + \|\tau w(\tau/2) - \tau T(\tau/2)B(u^*)T(\tau/2)u_0\|_{H^2} \qquad (3.3) \\ &=: S_1 + S_2. \end{aligned}$$

So, we have split the local error into a quadrature error and a remainder error term. The calculation

$$2\,\mathrm{Re}\big(\overline{u}(s)B(u(s))u(s)\big) = 2\,\mathrm{Re}\big(-i\mu\,|u(s)|^4\big) = 0$$

gives the identity

$$\partial_s B(u(s)) = -2i\mu\,\mathrm{Re}\Big(\overline{u}(s)\big(A + B(u(s))\big)u(s)\Big) = -2i\mu\,\mathrm{Re}\big(\overline{u}(s)Au(s)\big). \qquad (3.4)$$

Using this result, we infer

$$\begin{aligned} w'(s) &= -T(\tau - s)AB(u(s))T(s)u_0 \\ &\quad - 2i\mu T(\tau - s)\,\mathrm{Re}\big(\overline{u}(s)Au(s)\big)T(s)u_0 \qquad (3.5) \\ &\quad + T(\tau - s)B(u(s))T(s)Au_0. \end{aligned}$$

The algebra property of H^2 and H^4, and the unitarity of $T(\cdot)$ thus implies

$$\|w'(s)\|_{H^2} \le c\big(\|u(s)\|_{H^4}^2 \|u_0\|_{H^4} + \|u(s)\|_{H^2} \|u(s)\|_{H^4} \|u_0\|_{H^2} + \|u(s)\|_{H^2}^2 \|u_0\|_{H^4}\big).$$

Therefore,

$$\sup_{s\in[0,T]} \|w'(s)\|_{H^2} \le c\big(M_4^3 + M_2^2 M_4\big) \le cM_4^3.$$

Because the midpoint quadrature rule has order two (and hence also order one), we conclude from this calculation and Proposition 1.1 that

$$S_1 \le c \cdot \sup_{s\in[0,T]} \|w'(s)\|_{H^2} \cdot \tau^2 =: \widetilde{C}_{1,1}\tau^2, \tag{3.6}$$

with $\widetilde{C}_{1,1}$ only depending on M_4.

To deal with the summand S_2 in (3.3), we note that with the definitions of w and u^*, the remainder error term has the form

$$\begin{aligned} &\tau w(\tau/2) - \tau T(\tau/2)B(u^*)T(\tau/2)u_0 \\ &\quad = \tau T(\tau/2)\Big(B(u(\tau/2)) - B\big(T(\tau/2)u_0\big)\Big)T(\tau/2)u_0. \end{aligned} \tag{3.7}$$

We introduce the function $f : [0,T] \to H^2$ defined by

$$f(t) := \Big(B(u(t/2)) - B\big(T(t/2)u_0\big)\Big)T(t/2)u_0.$$

Identity (3.4) yields the derivative

$$\begin{aligned} 2f'(t) = -2i\mu&\Big(\mathrm{Re}\big(\overline{u}(t/2)Au(t/2)\big) - \mathrm{Re}\big((\overline{T(t/2)u_0)}AT(t/2)u_0\big)\Big)T(t/2)u_0 \\ &+ \Big(B(u(t/2)) - B\big(T(t/2)u_0\big)\Big)T(t/2)Au_0. \end{aligned} \tag{3.8}$$

We employ again the algebra property of H^2 and H^4 as well as the unitarity of $T(t/2)$ and obtain the inequalities

$$\begin{aligned} \|f'(t)\|_{H^2} &\le c\big(\|u(t/2)\|_{H^2}\|u(t/2)\|_{H^4}\|u_0\|_{H^2} + \|u_0\|_{H^2}\|u_0\|_{H^4}\|u_0\|_{H^2} \\ &\qquad + \|u(t/2)\|_{H^2}^2\|u_0\|_{H^4} + \|u_0\|_{H^2}^2\|u_0\|_{H^4}\big), \\ \sup_{t\in[0,T]} \|f'(t)\|_{H^2} &\le cM_2^2M_4 \le cM_4^3. \end{aligned} \tag{3.9}$$

Due to $f(0) = 0$, we have

$$f(\tau) = f(0) + \int_0^\tau f'(t)\,\mathrm{d}t = \int_0^\tau f'(t)\,\mathrm{d}t.$$

Hence, the formulas (3.7) and (3.9) lead to the estimate

$$S_2 = \|\tau T(\tau/2)f(\tau)\|_{H^2} \le \widetilde{C}_{1,2}\tau^2, \tag{3.10}$$

with $\widetilde{C}_{1,2}$ only depending on M_4.

2) *Bound on I_2*: By means of Lemma 2.1, we estimate the two summands of I_2 by

$$\left\| \int_0^\tau T(\tau - s)B(u(s)) \int_0^s T(s-\sigma)B(u(\sigma))u(\sigma)\,\mathrm{d}\sigma\,\mathrm{d}s \right\|_{H^2}$$

$$\leq c\int_0^\tau \|u(s)\|_{H^2}^2 \int_0^s \|u(\sigma)\|_{H^2}^3 \,\mathrm{d}\sigma\,\mathrm{d}s \leq c\frac{\tau^2}{2}M_4^5 =: \widetilde{C}_{2,1}\tau^2$$

and

$$\begin{aligned}
&\left\|\int_0^\tau (\tau-s)T(\tau/2)B(u^*)^2 e^{sB(u^*)}T(\tau/2)u_0\,\mathrm{d}s\right\|_{H^2}\\
&\quad\leq c\int_0^\tau (\tau-s)\,\|T(\tau/2)u_0\|_{H^2}^4\,\|u_0\|_{H^2}\,\mathrm{d}s\\
&\quad\leq c\frac{\tau^2}{2}\,\|u_0\|_{H^2}^5 \leq c\frac{\tau^2}{2}M_4^5 =: \widetilde{C}_{2,2}\tau^2,
\end{aligned}$$

using $\left\|e^{sB(u^*)}\right\|_{L^\infty} = 1$. With $\widetilde{C}_2 := \widetilde{C}_{2,1} + \widetilde{C}_{2,2}$, the summands of I_2 are together bounded by $\widetilde{C}_2\tau^2$, where $\widetilde{C}_2$ only depends on M_4.

We combine the two estimates above with (3.2), (3.3), (3.6) and (3.10) to finish the proof. □

3.2.2. Stability in the H^2-norm

Proof (of Lemma 3.4):
Let $z_0, w_0 \in H^2$ with $\|z_0\|_{H^2} \leq M$ and $\|w_0\|_{H^2} \leq M$. We first look at the initial value problem

$$\partial_t z(t) = -i\mu\,|z(t)|^2\,z(t), \qquad z(0) = z_0.$$

Its solution is $z(t) = \exp(-i\mu t\,|z_0|^2)z_0$ for all $t \geq 0$. We additionally set $w(t) := \exp(-i\mu t\,|w_0|^2)w_0$. The first and second derivatives of z are given by

$$\begin{aligned}
\partial_j z(t) &= -2i\mu t\exp(-i\mu t\,|z_0|^2)\,\mathrm{Re}(\overline{z_0}\partial_j z_0)z_0 + \exp(-i\mu t\,|z_0|^2)\partial_j z_0,\\
\partial_{jk} z(t) &= -4\mu^2 t^2\exp(-i\mu t\,|z_0|^2)\,\mathrm{Re}(\overline{z_0}\partial_k z_0)\,\mathrm{Re}(\overline{z_0}\partial_j z_0)z_0\\
&\quad - 2i\mu t\exp(-i\mu t\,|z_0|^2)\big(\mathrm{Re}(\overline{z_0}\partial_{jk} z_0) + \mathrm{Re}((\partial_j z_0)\partial_k\overline{z_0})\big)z_0\\
&\quad - 2i\mu t\exp(-i\mu t\,|z_0|^2)\,\mathrm{Re}(\overline{z_0}\partial_k z_0)\partial_j z_0\\
&\quad - 2i\mu t\exp(-i\mu t\,|z_0|^2)\,\mathrm{Re}(\overline{z_0}\partial_j z_0)\partial_k z_0\\
&\quad + \exp(-i\mu t\,|z_0|^2)\partial_{jk} z_0
\end{aligned}$$

for all $t \geq 0$ and $j, k \in \{1, \cdots, d\}$. Using the embeddings $H^2 \hookrightarrow L^\infty$ and $H^1 \hookrightarrow L^6$ as well as

$$\left\|\exp\big(-i\mu t\,|z_0|^2\big)\right\|_{L^\infty} = 1,$$

we deduce

$$\begin{aligned}
\|z(t)\|_{H^2} &\leq c\Big(\|z(t)\|_{L^2} + \sum_{j=1}^d \|\partial_j z(t)\|_{L^2} + \sum_{j,k=1}^d \|\partial_{jk} z(t)\|_{L^2}\Big)\\
&\leq c\Big(\|z_0\|_{L^2} + \big(t\,\|\nabla z_0\|_{L^2}\,\|z_0\|_{L^\infty}^2 + \|\nabla z_0\|_{L^2}\big)
\end{aligned}$$

$$+\Big(t^2\,\|\nabla z_0\|_{H^1}^2\,\|z_0\|_{H^1}\,\|z_0\|_{L^\infty}^2+t\,\big\|D^2z_0\big\|_{L^2}\,\|z_0\|_{L^\infty}^2 \\ +t\,\|\nabla z_0\|_{H^1}^2\,\|z_0\|_{H^1}+t\,\|\nabla z_0\|_{H^1}^2\,\|z_0\|_{H^1}+t\,\|\nabla z_0\|_{H^1}^2\,\|z_0\|_{H^1}+\big\|D^2z_0\big\|_{L^2}\Big)\Big)$$

for all $t\geq 0$, where D^2z_0 denotes the matrix of the second-order derivatives of z_0. This yields

$$\|z(t)\|_{H^2}\leq c\Big(\|z_0\|_{H^2}+t\,\|z_0\|_{H^2}^3+t^2\,\|z_0\|_{H^2}^5\Big) \tag{3.11}$$

for all $t\geq 0$. (If one simply applies Lemma 2.1 here, one obtains worse constants below.) We further compute

$$\begin{aligned}\partial_t z(t)-\partial_t w(t)&=-i\mu\big(|z_0|^2\,z(t)-|w_0|^2\,w(t)\big)\\&=-i\mu\big((z_0-w_0)\overline{z_0}\big)z(t)-i\mu\big(w_0(\overline{z_0}-\overline{w_0})\big)z(t)\\&\quad-i\mu w_0\overline{w_0}(z(t)-w(t)).\end{aligned}$$

Lemma 2.1 and estimate (3.11) then yield

$$\begin{aligned}&\|\partial_t z(t)-\partial_t w(t)\|_{H^2}\\&\quad\leq c\,\|z_0-w_0\|_{H^2}\,\|z_0\|_{H^2}\,\|z(t)\|_{H^2}+c\,\|w_0\|_{H^2}\,\|z_0-w_0\|_{H^2}\,\|z(t)\|_{H^2}\\&\qquad+c\,\|w_0\|_{H^2}^2\,\|z(t)-w(t)\|_{H^2}\\&\quad\leq c\big(\|z_0\|_{H^2}+\|w_0\|_{H^2}\big)\big(\|z_0\|_{H^2}+t\,\|z_0\|_{H^2}^3+t^2\,\|z_0\|_{H^2}^5\big)\,\|z_0-w_0\|_{H^2}\\&\qquad+c\,\|w_0\|_{H^2}^2\,\|z(t)-w(t)\|_{H^2}\,.\end{aligned} \tag{3.12}$$

Integrating from 0 to τ, we thus infer

$$\begin{aligned}\|z(\tau)-w(\tau)\|_{H^2}&=\Big\|z_0-w_0+\int_0^\tau\partial_t(z(t)-w(t))\,\mathrm{d}t\Big\|_{H^2}\\&\leq\|z_0-w_0\|_{H^2}+c\big(\|w_0\|_{H^2}+\|w_0\|_{H^2}\big)\big(\tau\,\|z_0\|_{H^2}\\&\qquad+\tfrac12\tau^2\,\|w_0\|_{H^2}^3+\tfrac13\tau^3\,\|z_0\|_{H^2}^5\big)\,\|z_0-w_0\|_{H^2}\\&\qquad+c\,\|w_0\|_{H^2}^2\int_0^\tau\|z(t)-w(t)\|_{H^2}\,\mathrm{d}t\\&\leq\|z_0-w_0\|_{H^2}+cM(\tau M+\tau^2M^3+\tau^3M^5)\,\|z_0-w_0\|_{H^2}\\&\qquad+cM^2\int_0^\tau\|z(t)-w(t)\|_{H^2}\,\mathrm{d}t.\end{aligned}$$

Gronwall's inequality now yields

$$\begin{aligned}\|z(\tau)-w(\tau)\|_{H^2}&\leq\big(1+cM^2\tau+(cM^2)^2\tfrac{\tau^2}{2}+(cM^2)^3\tfrac{\tau^3}{6}\big)\,\|z_0-w_0\|_{H^2}\,e^{cM^2\tau}\\&\leq e^{cM^2\tau}\,\|z_0-w_0\|_{H^2}\,.\end{aligned} \tag{3.13}$$

Let $u_0,v_0\in H^2$ with $\|u_0\|_{H^2}\leq M$ and $\|v_0\|_{H^2}\leq M$. Because $T(\tau/2)$ is unitary, we first have

$$\|T(\tau/2)u_0\|_{H^2}=\|u_0\|_{H^2}\leq M\qquad\text{and}\qquad\|T(\tau/2)v_0\|_{H^2}=\|v_0\|_{H^2}\leq M.$$

Therefore, estimate (3.13) leads to

$$\begin{aligned}
&\|\Psi_\tau(u_0) - \Psi_\tau(v_0)\|_{H^2} \\
&= \Big\| T(\tau/2) \exp\big(-i\mu\tau \,|T(\tau/2)u_0|^2\big) T(\tau/2)u_0 - T(\tau/2) \exp\big(-i\mu\tau \,|T(\tau/2)v_0|^2\big) T(\tau/2)v_0 \Big\|_{H^2} \\
&= \big\| \exp\big(-i\mu\tau \,|T(\tau/2)u_0|^2\big) T(\tau/2)u_0 - \exp\big(-i\mu\tau \,|T(\tau/2)v_0|^2\big) T(\tau/2)v_0 \big\|_{H^2} \\
&\leq e^{cM^2\tau} \|T(\tau/2)u_0 - T(\tau/2)v_0\|_{H^2} = e^{cM^2\tau} \|u_0 - v_0\|_{H^2} \,.
\end{aligned}$$

The claim follows with $C_2 := cM^2$. □

3.2.3. Boundedness of the numerical solution in the H^2-norm

Proof (of Lemma 3.6):
We denote by $u(s, y_0)$ the solution to problem (2.5) at time $s \geq 0$ with initial function $y_0 \in H^4$. Let $u_0 \in H^4$ and define

$$\tau_0 := \min\left\{ \frac{M_2}{T e^{TC_2} C_1}, T \right\}. \tag{3.14}$$

We prove part (b) and an even stronger version of part (a) with an induction argument. For all $\tau \in (0, \tau_0]$, $n \in \mathbb{N}_0$ with $n\tau \leq T$ and $k \in \{0, \dots, n\}$ we claim that

$$\big\| \Psi_\tau^{n-k}(u(k\tau, u_0)) - u(n\tau, u_0) \big\|_{H^2} \leq T e^{TC_2} C_1 \tau \tag{3.15}$$

with C_1 from Lemma 3.3 and C_2 from Lemma 3.4 (with $M := 2M_2$). We note that definition (3.14) and estimate (3.15) yield

$$\big\| \Psi_\tau^{n-k}(u(k\tau, u_0)) - u(n\tau, u_0) \big\|_{H^2} \leq M_2$$

for $\tau \in (0, \tau_0]$, so that the strong boundedness estimate

$$\big\| \Psi_\tau^{n-k}(u(k\tau, u_0)) \big\|_{H^2} \leq 2M_2 =: \widehat{C} \tag{3.16}$$

will follow from (3.15) with the triangle inequality.

We fix $\tau \in (0, \tau_0]$ and establish (3.15) by induction. The case $n = 0$ is trivial. Let the induction hypothesis

$$\big\| \Psi_\tau^{n-k}(u(k\tau, u_0)) - u(n\tau, u_0) \big\|_{H^2} \leq T e^{TC_2} C_1 \tau \leq M_2$$

hold true for some $n \in \mathbb{N}_0$ with $(n+1)\tau \leq T$ and all $k \in \{0, \dots, n\}$. Hence, (3.16) is valid for all $k \in \{0, \dots, n\}$. We now show (3.15) with n replaced by $n+1$. For $k = n+1$ estimate (3.15) is clear. Let $k \in \{0, \dots, n\}$. Estimate (3.16) for n gives a uniform constant C_2 for the following applications of Lemma 3.4, so that Lemma 3.4 and 3.3 imply via a telescopic sum that

$$\big\| \Psi_\tau^{n+1-k}(u(k\tau, u_0)) - u((n+1)\tau, u_0) \big\|_{H^2}$$

$$\leq \sum_{j=0}^{n-k} \left\| \Psi_\tau^{n-k-j}\big(\Psi_\tau\big(u((k+j)\tau, u_0)\big)\big) - \Psi_\tau^{n-k-j}\big(u((k+j+1)\tau, u_0)\big) \right\|_{H^2}$$
$$\leq \sum_{j=0}^{n-k} e^{(n-k-j)C_2\tau} \left\| \Psi_\tau\big(u((k+j)\tau, u_0)\big) - u\big(\tau, u((k+j)\tau, u_0)\big) \right\|_{H^2}$$
$$\leq \sum_{j=0}^{n-k} e^{(n-k-j)C_2\tau} C_1\tau^2 \leq \sum_{j=0}^{n-k} e^{TC_2} C_1\tau^2 \leq T e^{TC_2} C_1\tau.$$

To estimate the local errors with starting point $u(l\tau, u_0)$ in the second to the last line we use that for all $l \in \{0, \dots, n\}$ the constant C_1 from Lemma 3.3 only depends on

$$\sup_{t\in[0,T-l\tau]} \|u(t, u(l\tau, u_0))\|_{H^4} \leq M_4$$

and in particular not on l. □

3.3. The estimate in L^2

We first prove Lemma 3.7. Afterwards we show Lemma 3.8 and combine it with Lemma 3.7 and 3.6 to derive Theorem 3.1.

3.3.1. The local error in the L^2-norm

The proof of Lemma 3.7 is similar to the one of Lemma 3.3, but we need a Taylor expansion of second order instead of first order. We furthermore use the following non-standard *quadrature rule for two-dimensional simplexes*.

Lemma 3.9. *Let X be a Hilbert space. On the simplex*

$$S := \{(x, y) \in \mathbb{R}^2 \mid x, y \geq 0, x + y \leq 1\}$$

we choose the quadrature rule with the equally weighted nodes $\xi_1 := (0, 0)$, $\xi_2 := (1, 0)$, $\xi_3 := (0, 1)$ and $\xi_4 := (1/3, 1/3)$, i.e. for functions $f : S \to X$ we use the quadrature rule

$$Q(f) := \frac{1}{8}\big(f(0, 0) + f(1, 0) + f(0, 1) + f(1/3, 1/3)\big).$$

Let $\tau > 0$. Transforming this map to the shrunk, rotated and translated simplex

$$S_\tau := \{(x, y) \in \mathbb{R}^2 \mid 0 \leq y \leq x \leq \tau\}$$

gives for a function $\widetilde{f} : S_\tau \to X$ the quadrature rule

$$Q_\tau(\widetilde{f}) := \frac{\tau^2}{8}\big(\widetilde{f}(0, 0) + \widetilde{f}(\tau, 0) + \widetilde{f}(\tau, \tau) + \widetilde{f}(2\tau/3, \tau/3)\big).$$

These two quadrature rules have order two.

PROOF:
The simplex S is mapped bijectively onto the simplex S_τ by the linear transformation $(x, y) \mapsto (-y\tau + \tau, x\tau)$. Therefore, both quadrature rules have the same order. For an arbitrary affine $f : S \to X$, written as $f(x, y) = a_1 x + a_2 y + a_3$ with $a_1, a_2, a_3 \in X$, we compute

$$\begin{aligned}\int_S f(x,y)\,d(x,y) &= \int_0^1 \int_0^{1-x} (a_1 x + a_2 y + a_3)\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_0^1 \big(a_1 x(1-x) + \tfrac{1}{2}a_2(1-x)^2 + a_3(1-x)\big)\,\mathrm{d}x \\ &= \frac{1}{6}a_1 + \frac{1}{6}a_2 + \frac{1}{2}a_3, \\ Q(f) &= \frac{1}{8}\big(\tfrac{4}{3}a_1 + \tfrac{4}{3}a_2 + 4a_3\big) = \frac{1}{6}a_1 + \frac{1}{6}a_2 + \frac{1}{2}a_3.\end{aligned}$$

This shows that the two quadrature rules have order (at least) two. □

PROOF (OF LEMMA 3.7):
Let $u_0 \in H^4$ and $\tau \in (0, T]$. We use the Taylor expansion

$$e^{\tau x} = 1 + \tau x + \frac{\tau^2}{2}x^2 + \frac{1}{2}\int_0^\tau (\tau - s)^2 x^3 e^{sx}\,\mathrm{d}s$$

for $u^{**} = \exp\big(\tau B(u^*)\big)u^*$ and obtain

$$u^{**} = u^* + \tau B(u^*)u^* + \frac{\tau^2}{2}B(u^*)^2 u^* + \frac{1}{2}\int_0^\tau (\tau - s)^2 B(u^*)^3 e^{sB(u^*)}u^*\,\mathrm{d}s. \tag{3.17}$$

Recall that $\Psi_\tau(u_0) = T(\tau/2)u^{**}$ and $u^* = T(\tau/2)u_0$, see definition (2.10). We apply $T(\tau/2)$ to (3.17) and insert $u^* = T(\tau/2)u_0$ thrice, arriving at

$$\begin{aligned}\Psi_\tau(u_0) = T(\tau)u_0 &+ \tau T(\tau/2)B(u^*)T(\tau/2)u_0 + \frac{\tau^2}{2}T(\tau/2)B(u^*)^2 T(\tau/2)u_0 \\ &+ \frac{1}{2}\int_0^\tau (\tau - s)^2 T(\tau/2)B(u^*)^3 e^{sB(u^*)}u^*\,\mathrm{d}s.\end{aligned}$$

Subtracting this identity from the representation (3.1) for $u(\tau)$, we infer

$$\begin{aligned}&u(\tau) - \Psi_\tau(u_0) \\ &\quad= \left(\int_0^\tau T(\tau - s)B(u(s))T(s)u_0\,\mathrm{d}s - \tau T(\tau/2)B(u^*)T(\tau/2)u_0\right) \\ &\qquad+ \left(\int_0^\tau T(\tau - s)B(u(s))\int_0^s T(s - \sigma)B(u(\sigma))u(\sigma)\,\mathrm{d}\sigma\,\mathrm{d}s\right. \\ &\qquad\qquad\left. - \frac{\tau^2}{2}T(\tau/2)B(u^*)^2 T(\tau/2)u_0\right)\end{aligned} \tag{3.18}$$

$$
\begin{aligned}
&- \frac{1}{2}\int_0^\tau (\tau-s)^2 T(\tau/2)B(u^*)^3 e^{sB(u^*)}u^*\,\mathrm{d}s \\
&=: I_1 + I_2 + I_3.
\end{aligned}
$$

1) *Bound on* I_1: We first introduce the function $w : [0,T] \to L^2$ by

$$
w(s) := T(\tau-s)B(u(s))T(s)u_0,
$$

see Section 3.2. With the midpoint quadrature rule we split I_1 into a quadrature error and a remainder error term, which yields

$$
\begin{aligned}
\|I_1\|_{L^2} &\le \left\| \int_0^\tau T(\tau-s)B(u(s))T(s)u_0\,\mathrm{d}s - \tau w(\tau/2) \right\|_{L^2} \\
&\quad + \|\tau w(\tau/2) - \tau T(\tau/2)B(u^*)T(\tau/2)u_0\|_{L^2} \\
&=: S_1 + S_2.
\end{aligned} \tag{3.19}
$$

In (3.5) we have seen

$$
\begin{aligned}
w'(s) = &-T(\tau-s)AB(u(s))T(s)u_0 \\
&- 2i\mu T(\tau-s)\,\mathrm{Re}\big(\overline{u}(s)Au(s)\big)T(s)u_0 \\
&+ T(\tau-s)B(u(s))AT(s)u_0.
\end{aligned}
$$

By differentiating, reordering the terms and using the identities (3.4) and (2.5), we conclude

$$
\begin{aligned}
w''(s) = &\, T(\tau-s)A^2B(u(s))T(s)u_0 - 2T(\tau-s)AB(u(s))T(s)Au_0 \\
&+ T(\tau-s)B(u(s))T(s)A^2u_0 \\
&+ 4i\mu T(\tau-s)A\Big(\mathrm{Re}\big(\overline{u}(s)Au(s)\big)T(s)u_0\Big) \\
&- 2i\mu T(\tau-s)\,|Au(s)|^2\,T(s)u_0 \\
&- 2i\mu T(\tau-s)\,\mathrm{Re}\big(B(u(s))u(s)\overline{Au(s)}\big)T(s)u_0 \\
&- 2i\mu T(\tau-s)\,\mathrm{Re}\big(\overline{u}(s)A^2u(s)\big)T(s)u_0 \\
&- 2i\mu T(\tau-s)\,\mathrm{Re}\big(\overline{u}(s)AB(u(s))u(s)\big)T(s)u_0 \\
&- 4i\mu T(\tau-s)\,\mathrm{Re}\big(\overline{u}(s)Au(s)\big)T(s)Au_0.
\end{aligned}
$$

We again employ that $T(\cdot)$ is unitary, that H^2 and H^4 are algebras and the Sobolev embedding $H^2 \hookrightarrow L^\infty$, and estimate

$$
\begin{aligned}
\|w''(s)\|_{L^2} \le c\big(&\|u(s)\|_{H^4}^2\,\|u_0\|_{H^4} + \|u(s)\|_{H^2}^2\,\|u_0\|_{H^4} \\
&+ \|u(s)\|_{H^2}^2\,\|u_0\|_{H^4} + \|u(s)\|_{H^2}\,\|u(s)\|_{H^4}\,\|u_0\|_{H^2} \\
&+ \|u(s)\|_{H^2}\,\|u(s)\|_{H^4}\,\|u_0\|_{H^2} + \|u(s)\|_{H^2}^4\,\|u_0\|_{H^2}
\end{aligned}
$$

$$
\begin{aligned}
&+ \|u(s)\|_{H^2} \|u(s)\|_{H^4} \|u_0\|_{H^2} + \|u(s)\|_{H^2}^4 \|u_0\|_{H^2} \\
&+ \|u(s)\|_{H^2}^2 \|u_0\|_{H^4}\big).
\end{aligned}
$$

As a result,

$$
\sup_{s\in[0,T]} \|w''(s)\|_{L^2} \le c\big(M_4^3 + M_2^2 M_4 + M_2^5\big) \le c(M_4^3 + M_4^5).
$$

Since the midpoint quadrature rule has order two, we conclude from Proposition 1.1 that

$$
S_1 \le c \cdot \sup_{s\in[0,T]} \|w''(s)\|_{L^2} \cdot \tau^3 =: \widetilde{C}_{2,1}\tau^3, \tag{3.20}
$$

with $\widetilde{C}_{3,1}$ only depending on M_4.

For the treatment of S_2 from (3.19), as in Section 3.2, we define the function $f : [0,T] \to L^2$ by

$$
f(t) := \Big(B(u(t/2)) - B\big(T(t/2)u_0\big)\Big)T(t/2)u_0. \tag{3.21}
$$

We recall formula (3.8),

$$
\begin{aligned}
2f'(t) = &-2i\mu\Big(\mathrm{Re}\big(\overline{u}(t/2)Au(t/2)\big) - \mathrm{Re}\big(\overline{(T(t/2)u_0)}AT(t/2)u_0\big)\Big)T(t/2)u_0 \\
&+ \Big(B(u(t/2)) - B\big(T(t/2)u_0\big)\Big)T(t/2)Au_0.
\end{aligned}
$$

By means of the identities (3.4) and (2.5) we further compute

$$
\begin{aligned}
4f''(t) = &-2i\mu\Big(|Au(t/2)|^2 - |AT(t/2)u_0|^2\Big)T(t/2)u_0 \\
&- 2i\mu\Big(\mathrm{Re}\big(\overline{u}(t/2)A^2u(t/2)\big) - \mathrm{Re}\big(\overline{T(t/2)u_0}A^2T(t/2)u_0\big)\Big)T(t/2)u_0 \\
&- 2i\mu\Big(\mathrm{Re}\big(\overline{B(u(t/2))u(t/2)}Au(t/2)\big)\Big)T(t/2)u_0 \\
&- 2i\mu\Big(\mathrm{Re}\big(\overline{u}(t/2)A\big(B(u(t/2))u(t/2)\big)\big)\Big)T(t/2)u_0 \\
&- 4i\mu\Big(\mathrm{Re}\big(\overline{u}(t/2)Au(t/2)\big) - \mathrm{Re}\big(\overline{T(t/2)u_0}AT(t/2)u_0\big)\Big)T(t/2)Au_0 \\
&+ \Big(B(u(t/2)) - B\big(T(t/2)u_0\big)\Big)T(t/2)A^2u_0.
\end{aligned}
$$

Using Lemma 2.1 and the unitarity of $T(t/2)$, we conclude

$$
\begin{aligned}
\|f''(t)\|_{L^2} \le c\big(&\|u(t/2)\|_{H^2} \|u(t/2)\|_{H^4} \|u_0\|_{H^2} + \|u_0\|_{H^2}^2 \|u_0\|_{H^4} \\
&+ \|u(t/2)\|_{H^2} \|u(t/2)\|_{H^4} \|u_0\|_{H^2} + \|u_0\|_{H^2}^2 \|u_0\|_{H^4} \\
&+ \|u(t/2)\|_{H^2}^4 \|u_0\|_{H^2} + \|u(t/2)\|_{H^2}^4 \|u_0\|_{H^2} \\
&+ \|u(t/2)\|_{H^2}^2 \|u_0\|_{H^4} + \|u_0\|_{H^2}^2 \|u_0\|_{H^4} \\
&+ \|u(t/2)\|_{H^2}^2 \|u_0\|_{H^4} + \|u_0\|_{H^2}^2 \|u_0\|_{H^4}\big), \\
\sup_{t\in[0,T]} \|f''(t)\|_{L^2} \le c\big(&M_2^2 M_4 + M_2^5\big) \le c\big(M_4^3 + M_4^5\big).
\end{aligned} \tag{3.22}
$$

With $f(0) = 0$ and $f'(0) = 0$ we get

$$f(\tau) = f(0) + f'(0) + \int_0^\tau (\tau - t) f''(t)\,\mathrm{d}t = \int_0^\tau (\tau - t) f''(t)\,\mathrm{d}t.$$

The inequality (3.22) thus implies

$$S_2 = \|\tau T(\tau/2) f(\tau)\|_{L^2} \le \widetilde{C}_{3,2}\tau^3, \tag{3.23}$$

with $\widetilde{C}_{3,2}$ only depending on M_4.

2) *Bound on I_2*: We rewrite

$$\begin{aligned}
&\int_0^\tau T(\tau - s)B(u(s)) \int_0^s T(s - \sigma)B(u(\sigma))u(\sigma)\,\mathrm{d}\sigma\,\mathrm{d}s \\
&\quad = \int_0^\tau \int_0^s T(\tau - s)B(u(s))T(s - \sigma)B(u(\sigma))u(\sigma)\,\mathrm{d}\sigma\,\mathrm{d}s.
\end{aligned}$$

We look at the function $v : [0, T] \times [0, T] \to L^2$ given by

$$v(s, \sigma) := T(\tau - s)B(u(s))T(s - \sigma)B(u(\sigma))u(\sigma). \tag{3.24}$$

As we did with the summand I_1, we split the term I_2 into a quadrature error and a remainder error term, namely

$$\begin{aligned}
\|I_2\|_{L^2} \le &\Bigg\| \int_0^\tau \int_0^s T(\tau - s)B(u(s))T(s - \sigma)B(u(\sigma))u(\sigma)\,\mathrm{d}\sigma\,\mathrm{d}s \\
&\qquad - \frac{\tau^2}{8}\big(v(0,0) + v(\tau,0) + v(\tau,\tau) + v(2\tau/3,\tau/3)\big)\Bigg\|_{L^2} \\
&\quad + \Bigg\| \frac{\tau^2}{8}\big(v(0,0) + v(\tau,0) + v(\tau,\tau) + v(2\tau/3,\tau/3)\big) \\
&\qquad - \frac{\tau^2}{2} T(\tau/2)B(u^*)^2 T(\tau/2)u_0\Bigg\|_{L^2} \\
=: &\, R_1 + R_2.
\end{aligned} \tag{3.25}$$

Using once more identity (3.4) yields

$$\begin{aligned}
\partial_s v(s,\sigma) = &-T(\tau - s)AB(u(s))T(s - \sigma)B(u(\sigma))u(\sigma) \\
&- 2i\mu T(\tau - s)\,\mathrm{Re}\big(\overline{u}(s)Au(s)\big)T(s - \sigma)B(u(\sigma))u(\sigma) \\
&+ T(\tau - s)B(u(s))T(s - \sigma)AB(u(\sigma))u(\sigma), \\
\partial_\sigma v(s,\sigma) = &-T(\tau - s)B(u(s))T(s - \sigma)AB(u(\sigma))u(\sigma) \\
&- 2i\mu T(\tau - s)B(u(s))T(s - \sigma)\,\mathrm{Re}\big(\overline{u}(\sigma)Au(\sigma)\big)u(\sigma) \\
&+ T(\tau - s)B(u(s))T(s - \sigma)B(u(\sigma))\big(Au(\sigma) + B(u(\sigma))u(\sigma)\big).
\end{aligned}$$

Estimating as above, we derive

$$\begin{aligned}\|\partial_s v(s,\sigma)\|_{L^2} &\le c\big(\|u(s)\|_{H^2}^2\,\|u(\sigma)\|_{H^2}^3 + \|u(s)\|_{H^2}^2\,\|u(\sigma)\|_{H^2}^3 \\ &\qquad + \|u(s)\|_{H^2}\,\|u(\sigma)\|_{H^2}^3\big),\\ \|\partial_\sigma v(s,\sigma)\|_{L^2} &\le c\Big(\|u(s)\|_{H^2}^2\,\|u(\sigma)\|_{H^2}^3 + \|u(s)\|_{H^2}^2\,\|u(\sigma)\|_{H^2}^3 \\ &\qquad + \|u(s)\|_{H^2}^2\,\|u(\sigma)\|_{H^2}^2\,\big(\|u(\sigma)\|_{H^2} + \|u(\sigma)\|_{H^2}^2\,\|u(\sigma)\|_{L^2}\big)\Big).\end{aligned}$$

So, we have

$$\begin{aligned}\sup_{(s,\sigma)\in[0,T]\times[0,T]} \|\partial_s v(s,\sigma)\|_{L^2} &\le cM_2^5\\ \sup_{(s,\sigma)\in[0,T]\times[0,T]} \|\partial_\sigma v(s,\sigma)\|_{L^2} &\le c\big(M_2^5 + M_0M_2^6\big) \le c\big(M_2^5 + M_2^7\big).\end{aligned}$$

Lemma 3.9, then implies the bound

$$R_1 \le c\cdot \max_{(s,\sigma)\in[0,T]\times[0,T]} \left|\begin{pmatrix} \|\partial_s v(s,\sigma)\|_{L^2} \\ \|\partial_\sigma w(s,\sigma)\|_{L^2}\end{pmatrix}\right|_2 \cdot \tau^3 =: \widetilde{C}_{4,1}\tau^3, \tag{3.26}$$

with $\widetilde{C}_{4,1}$ only depending on M_2.

To control the remainder error term R_2 in (3.25), we notice

$$\begin{aligned} v(0,0) &= T(\tau)B(u_0)^2u_0,\\ v(\tau,0) &= B(u(\tau))T(\tau)B(u_0)u_0,\\ v(\tau,\tau) &= B(u(\tau))^2u(\tau) \qquad \text{and}\\ v(2\tau/3,\tau/3) &= T(\tau/3)B(u(2\tau/3))T(\tau/3)B(u(\tau/3))u(\tau/3).\end{aligned}$$

Hence, R_2 becomes

$$\begin{aligned}R_2 = \Big\| \frac{\tau^2}{8}\Big(&T(\tau)B(u_0)^2u_0 + B(u(\tau))T(\tau)B(u_0)u_0 + B(u(\tau))^2u(\tau)\\ &+ T(\tau/3)B(u(2\tau/3))T(\tau/3)B(u(\tau/3))u(\tau/3)\Big)\\ &- \frac{\tau^2}{2}T(\tau/2)B(u^*)^2T(\tau/2)u_0\Big\|_{L^2}.\end{aligned}$$

We introduce the functions $g_1, g_2, g_3, g_4, h, g : [0,T] \to L^2$ by

$$\begin{aligned} g_1(t) &:= T(t)B(u_0)^2u_0,\\ g_2(t) &:= B(u(t))T(t)B(u_0)u_0,\\ g_3(t) &:= B(u(t))^2u(t),\\ g_4(t) &:= T(t/3)B(u(2t/3))T(t/3)B(u(t/3))u(t/3),\end{aligned}$$

$$
\begin{aligned}
h(t) &:= T(t/2)B(T(t/2)u_0)^2T(t/2)u_0,\\
g &:= g_1 + g_2 + g_3 + g_4 - 4h.
\end{aligned}
$$

Identity (3.4) then yields the derivatives

$$
\begin{aligned}
g_1'(t) &= AT(t)B(u_0)^2u_0,\\
g_2'(t) &= -2i\mu\,\mathrm{Re}\big(\overline{u(t)}Au(t)\big)T(t)B(u_0)u_0 + B(u(t))AT(t)B(u_0)u_0,\\
g_3'(t) &= -4i\mu\,\mathrm{Re}\big(\overline{u(t)}Au(t)\big)B(u(t))u(t) + B(u(t))^2\big(Au(t) + B(u(t))u(t)\big),\\
3g_4'(t) &= AT(t/3)B(u(2t/3))T(t/3)B(u(t/3))u(t/3)\\
&\quad - 4i\mu T(t/3)\,\mathrm{Re}\big(\overline{u}(2t/3)Au(2t/3)\big)T(t/3)B(u(t/3))u(t/3)\\
&\quad + T(t/3)B(u(2t/3))AT(t/3)B(u(t/3))u(t/3)\\
&\quad - 2i\mu T(t/3)B(u(2t/3))T(t/3)\,\mathrm{Re}\big(\overline{u}(t/3)Au(t/3)\big)u(t/3)\\
&\quad + T(t/3)B(u(2t/3))T(t/3)B(u(t/3))\big(Au(t/3) + B(u(t/3))u(t/3)\big),\\
2h'(t) &= AT(t/2)B(T(t/2)u_0)^2T(t/2)u_0\\
&\quad - 4i\mu\,\mathrm{Re}\big((\overline{T(t/2)u_0})AT(t/2)u_0\big)B(T(t/2)u_0)T(t/2)u_0\\
&\quad + T(t/2)B(T(t/2)u_0)^2T(t/2)Au_0.
\end{aligned}
$$

As before these derivatives can bounded by

$$
\begin{aligned}
\|g_1'(t)\|_{L^2} &\le c\,\|u_0\|_{H^2}^5\,,\\
\|g_2'(t)\|_{L^2} &\le c\big(\|u(t)\|_{H^2}^2\,\|u_0\|_{H^2}^3 + \|u(t)\|_{H^2}^2\,\|u_0\|_{H^2}^3\big),\\
\|g_3'(t)\|_{L^2} &\le c\Big(\|u(t)\|_{H^2}^5 + \|u(t)\|_{H^2}^4\,\big(\|u(t)\|_{H^2} + \|u(t)\|_{H^2}^2\,\|u(t)\|_{L^2}\big)\Big),\\
\|g_4'(t)\|_{L^2} &\le c\Big(\|u(2t/3)\|_{H^2}^2\,\|u(t/3)\|_{H^2}^3 + \|u(2t/3)\|_{H^2}^2\,\|u(t/3)\|_{H^2}^3\\
&\qquad + \|u(2t/3)\|_{H^2}^2\,\|u(t/3)\|_{H^2}^3 + \|u(2t/3)\|_{H^2}^2\,\|u(t/3)\|_{H^2}^3\\
&\qquad + \|u(2t/3)\|_{H^2}^2\,\|u(t/3)\|_{H^2}^2\,\big(\|u(t/3)\|_{H^2} + \|u(t/3)\|_{H^2}^2\,\|u(t/3)\|_{L^2}\big)\Big),\\
\|h'(t)\|_{L^2} &\le c\big(\|u_0\|_{H^2}^5 + \|u_0\|_{H^2}^5 + \|u_0\|_{H^2}^5\big).
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
\sup_{t\in[0,T]}\|g_1'(t)\|_{L^2} &\le cM_2^5,\\
\sup_{t\in[0,T]}\|g_2'(t)\|_{L^2} &\le cM_2^5,\\
\sup_{t\in[0,T]}\|g_3'(t)\|_{L^2} &\le c\big(M_2^5 + M_2^7\big),\\
\sup_{t\in[0,T]}\|g_4'(t)\|_{L^2} &\le c(M_2^5 + M_2^7\big),\\
\sup_{t\in[0,T]}\|h'(t)\|_{L^2} &\le cM_2^5.
\end{aligned}
$$

Because of

$$g(0) = g_1(0) + g_2(0) + g_3(0) + g_4(0) - 4h(0) = 0,$$

g can be expressed by

$$g(\tau) = g(0) + \int_0^\tau g'(t)\,\mathrm{d}t = \int_0^\tau \big(g_1'(t) + g_2'(t) + g_3'(t) + g_4'(t) - 4h'(t)\big)\,\mathrm{d}t.$$

The bounds for the derivatives thus give

$$R_2 = \left\| \frac{\tau^2}{8} g(\tau) \right\|_{L^2} \leq \widetilde{C}_{4,2}\tau^3, \tag{3.27}$$

with $\widetilde{C}_{4,2}$ only depending on M_4.

3) *Bound on* I_3: Using $\left\|e^{sB(u^*)}\right\|_{L^\infty} = 1$, we estimate

$$\begin{aligned}
&\left\| \frac{1}{2} \int_0^\tau (\tau - s)^2 T(\tau/2) B(u^*)^3 e^{sB(u^*)} u^* \,\mathrm{d}s \right\|_{L^2} \\
&\quad \leq c \int_0^\tau (\tau - s)^2 \left\| B\big(T(\tau/2)u_0\big) \right\|_{H^2}^3 \|T(\tau/2)u_0\|_{L^2} \,\mathrm{d}s \\
&\quad \leq c\frac{\tau^3}{3} \|u_0\|_{H^2}^6 \|u_0\|_{L^2} \leq c\frac{\tau^3}{3} M_4^7 =: \widetilde{C}_5\tau^3,
\end{aligned}$$

with $\widetilde{C}_5$ only depending on M_4.

The claim now follows by combing the above estimate with (3.18), (3.19), (3.20), (3.23), (3.25), (3.26) and (3.27). □

3.3.2. H^2-conditional stability in the L^2-norm

Proof (of Lemma 3.8):
Let $u_0, v_0 \in H^2$ with $\|u_0\|_{H^2} \leq M$ and $\|v_0\|_{H^2} \leq M$. For $z_0, w_0 \in H^2$, we look at the solutions of the initial value problems

$$\begin{aligned}
\partial_t z(t) &= -i\mu \left|z(t)\right|^2 z(t), \qquad z(0) = z_0, \\
\partial_t w(t) &= -i\mu \left|w(t)\right|^2 w(t), \qquad w(0) = w_0.
\end{aligned}$$

As estimate (3.12) in the proof of Lemma 3.4, we derive

$$\begin{aligned}
\|\partial_t z(t) - \partial_t w(t)\|_{L^2} \leq c\big(\|z_0\|_{H^2} + \|w_0\|_{H^2}\big)\big(\|z_0\|_{H^2} + t\,\|z_0\|_{H^2}^3 \\
+ t^2\,\|z_0\|_{H^2}^5\big)\,\|z_0 - w_0\|_{L^2} + c\,\|w_0\|_{H^2}^2\,\|z(t) - w(t)\|_{L^2}\,.
\end{aligned}$$

From this fact we conclude with Gronwall's inequality that

$$\|z(t) - w(t)\|_{L^2} \leq e^{C_4\tau}\,\|z_0 - w_0\|_{L^2}$$

for a constant C_4 only depending on M, cf. (3.13). As in the proof of Lemma 3.4, we then arrive at

$$\|\Psi_\tau(u_0) - \Psi_\tau(v_0)\|_{L^2} \leq e^{C_4\tau}\,\|u_0 - v_0\|_{L^2}\,,$$

which is the desired estimate. □

3.3.3. Convergence in the L^2-norm

Proof (of Theorem 3.1):
Let $u_0 \in H^4$. Let $\tau \in (0, \tau_0]$ with $\tau_0 > 0$ from Lemma 3.6 and $n \in \mathbb{N}$ with $n\tau \leq T$. We have

$$u(n\tau) - \Psi_\tau^n(u_0) = \sum_{k=0}^{n-1} \Psi_\tau^k\big(u((n-k)\tau)\big) - \Psi_\tau^{k+1}\big(u((n-k-1)\tau)\big).$$

In view of Lemma 3.6, the expressions $\Psi_\tau^l\big(u((n-l)\tau)\big)$ with $l \in \{0, \dots, n\}$ are bounded in H^2 by a constant $\widehat{C}$ that only depends on M_2 (and in particular not on n or τ). Thus, Lemma 3.8 can iteratively be applied with $M := \widehat{C}$ to all summands appearing in the second line of the following calculation. Together with the local error bound in Lemma 3.7 we derive

$$\begin{aligned}
&\|u(n\tau) - \Psi_\tau^n(u_0)\|_{L^2} \\
&\quad\leq \sum_{k=0}^{n-1} \big\|\Psi_\tau^k\big(u((n-k)\tau)\big) - \Psi_\tau^{k+1}\big(u((n-k-1)\tau)\big)\big\|_{L^2} \\
&\quad\leq \sum_{k=0}^{n-1} e^{kC_4\tau} \big\|u\big(\tau, u((n-k-1)\tau)\big) - \Psi_\tau\big(u((n-k-1)\tau)\big)\big\|_{L^2} \\
&\quad\leq \sum_{k=0}^{n-1} e^{kC_4\tau} C_3\tau^3 \leq \sum_{k=0}^{n-1} e^{TC_4} C_3\tau^3 \leq T e^{TC_4} C_3\tau^2.
\end{aligned}$$

As in the analogous situation in the proof of Lemma 3.6 we use that for all $l \in \{0, \dots, n-1\}$ the constant C_3 from Lemma 3.7 for the local error with initial value $u(l\tau, u_0)$ only depends on

$$\sup_{t\in[0,T-l\tau]} \|u(t, u(l\tau, u_0))\|_{H^4} \leq M_4$$

and not on l. This completes the proof of Theorem 3.1. □

4. Convergence of the Strang splitting for initial functions in $H^{2+2\theta}$

Our aim is to show that the Strang splitting also converges if the initial function has a lower regularity than H^4. In this chapter we deal with the situation that the initial function is in $H^{2+2\theta}$ for $\theta \in (0,1)$. In Section 4.1 we state that the Strang splitting still converges but suffers from an order reduction that reduces the convergence order to $1+\theta$. We add some auxiliary results on the splitting scheme and the strategy of the proof, which are very similar to the ones of the H^4-situation in Chapter 3. The main difference in the proof is that we invoke interpolation estimates to cope with the reduced regularity. The details of the proof are presented in the Sections 4.2 and 4.3, separated according to arguments in H^2 and in L^2.

4.1. The theorem for initial functions in $H^{2+2\theta}$

The main result of this chapter is the following *fractional convergence theorem* for the Strang splitting.

Theorem 4.1. *For each $\theta \in (0,1)$ and $u_0 \in H^{2+2\theta}$, there exists a bound $\tau_0 > 0$ on the time step size such that we have*

$$\|u(n\tau) - \Psi_\tau^n(u_0)\|_{L^2} \leq C\tau^{1+\theta}$$

for all $\tau \in (0,\tau_0]$ and $n \in \mathbb{N}$ with $n\tau \leq T$ with a constant $C \geq 0$ that depends only on u_0 and T. More precisely, C depends only on T and $M_{2+2\theta}$.

The number $\tau_0 = \tau_0(\theta, T, M_2)$ is given in Lemma 4.3. It is possible to get rid of the dependency of τ_0 on θ, see Remark 5.9. We first show that the local error in H^2 is of order $1+\theta$.

Lemma 4.2. *For all $\theta \in (0,1)$, $u_0 \in H^{2+2\theta}$ and $\tau \in (0,T]$, we have*

$$\|u(\tau) - \Psi_\tau(u_0)\|_{H^2} \leq C_1\tau^{1+\theta},$$

with a constant $C_1 \geq 0$ depending only on T and $M_{2+2\theta}$.

As in Chapter 3, the precise form of the constant in the estimate is important since its n-th power enters in the proof of the main result. Also as in Chapter 3, our numerical solutions are strongly bounded in H^2.

Lemma 4.3. *Let $\theta \in (0,1)$ and $u_0 \in H^{2+2\theta}$. Then there exists a bound $\tau_0 > 0$ on the time step size, which is given by*

$$\tau_0 := \min\left\{ \left(\frac{M_2}{T e^{TC_2} C_1} \right)^{1/\theta}, T \right\},$$

with C_1 from Lemma 4.2 and C_2 from Lemma 3.4, such that the following two statements hold true.

(a) For all $\tau \in (0, \tau_0]$ and $n \in \mathbb{N}$ with $n\tau \leq T$, we have

$$\|\Psi_\tau^n(u_0) - u(n\tau)\|_{H^2} \leq C\tau^\theta,$$

with a constant $C \geq 0$ depending only on T and $M_{2+2\theta}$, i.e. the Strang splitting converges in H^2 with order θ.

(b) Ψ_τ is strongly bounded for (2.5) in H^2 for initial functions in $H^{2+2\theta}$, i.e. there exists a constant $\widehat{C} \geq 0$, only depending on T and M_2, such that $\left\|\Psi_\tau^{n-k}(u(k\tau))\right\|_{H^2} \leq \widehat{C}$ for all $\tau \in (0, \tau_0]$ and $n \in \mathbb{N}$ with $n\tau \leq T$ and $k \in \{0, \ldots, n\}$. In particular, the numerical solution is bounded in H^2 (choose $k = 0$).

The above lemmas are proved in Section 4.2. As in Chapter 3, we see in the next lemma that the order of the local error in L^2 is one higher than the one in H^2, namely $2 + \theta$ instead of $1 + \theta$.

Lemma 4.4. *For all $\theta \in (0,1)$, $u_0 \in H^{2+2\theta}$ and $\tau \in (0, T]$, we have*

$$\|u(\tau) - \Psi_\tau(u_0)\|_{L^2} \leq C_3 \tau^{2+\theta},$$

with a constant $C_3 \geq 0$ depending only on T and $M_{2+2\theta}$.

Together with the H^2-conditional stability from Lemma 3.8 we will obtain the desired result with Lady Windermere's fan.

4.2. The estimate in H^2

We prove Lemma 4.2 and combine it with Lemma 3.4 to conclude Lemma 4.3.

4.2.1. The local error in the H^2-norm

Proof (of Lemma 4.2):
Let $\theta > 0$, $u_0 \in H^{2+2\theta}$ and $\tau > 0$. We start our investigations with the representation (3.2) for $u(\tau) - \Psi_\tau(u_0)$, given by

$$\begin{aligned} u(\tau) - \Psi_\tau(u_0) &= \left(\int_0^\tau T(\tau - s) B(u(s)) T(s) u_0 \, \mathrm{d}s - \tau T(\tau/2) B(u^*) T(\tau/2) u_0 \right) \\ &\quad + \left(\int_0^\tau T(\tau - s) B(u(s)) \int_0^s T(s - \sigma) B(u(\sigma)) u(\sigma) \, \mathrm{d}\sigma \, \mathrm{d}s \right. \\ &\quad \left. - \int_0^\tau (\tau - s) T(\tau/2) B(u^*)^2 e^{sB(u^*)} T(\tau/2) u_0 \, \mathrm{d}s \right) \qquad (4.1) \\ &=: I_1 + I_2. \end{aligned}$$

1) *Bound on* I_1: As in Section 3.2, we use the function

$$w : [0, T] \to H^2; \qquad w(s) := T(\tau - s) B(u(s)) T(s) u_0,$$

and the estimate

$$\begin{aligned} \|I_1\|_{H^2} &\le \left\| \int_0^\tau w(s) \, \mathrm{d}s - \tau w(\tau/2) \right\|_{H^2} + \|\tau w(\tau/2) - \tau T(\tau/2) B(u^*) T(\tau/2) u_0\|_{H^2} \qquad (4.2) \\ &=: S_1 + S_2, \end{aligned}$$

see (3.3). For each $y \in H^{2+2\theta}$, the maps $t \mapsto T(t)y$ and $t \mapsto u(t, y)$ are θ-Hölder continuous in H^2 on $[0, T]$ by Lemma 2.3. From

$$\begin{aligned} w(s_1) - w(s_2) &= T(\tau - s_1) B(u(s_1)) \big(T(s_1) u_0 - T(s_2) u_0 \big) \\ &\quad - T(\tau - s_1) \big(B(u(s_1)) - B(u(s_2)) \big) T(s_2) u_0 \\ &\quad + \big(T(\tau - s_1) - T(\tau - s_2) \big) B(u(s_2)) T(s_2) u_0 \end{aligned}$$

we then deduce with the unitarity of $T(\cdot)$ that

$$\begin{aligned} \|w(s_1) - w(s_2)\|_{H^2} \le c \Big(& \|u(s_1)\|_{H^2}^2 \, \|u_0\|_{H^{2+2\theta}} \\ & + C(M_{2+2\theta}, T) \big(\|u(s_1)\|_{H^2} + \|u(s_2)\|_{H^2} \big) \, \|u_0\|_{H^2} \\ & + \|u(s_2)\|_{H^{2+2\theta}}^2 \, \|u_0\|_{H^{2+2\theta}} \Big) \cdot |s_1 - s_2|^\theta \end{aligned}$$

for all $s_1, s_2 \in [0, T)$. Here, we also took the algebra property from Lemma 2.1 into account. By this inequality, w belongs to $C^{0,\theta}([0, T], H^2)$ and

$$\begin{aligned} \|w(s_1) - w(s_2)\|_{H^2} &\le c \big(M_2^2 M_{2+2\theta} + C(M_{2+2\theta}, T) M_2^2 + M_{2+2\theta}^3 \big) \, |s_1 - s_2|^\theta \\ &\le c \big(M_{2+2\theta}^3 + C(M_{2+2\theta}, T) M_{2+2\theta}^2 \big) \, |s_1 - s_2|^\theta . \end{aligned}$$

4. Convergence of the Strang splitting for initial functions in $H^{2+2\theta}$

The space $C^{0,\theta}([0,T],H^2)$ of Hölder continuous functions is the real interpolation space

$$\big(C([0,T],H^2),C^1([0,T],H^2)\big)_{\theta,\infty}.$$

This can be proved as in the scalar-valued case, see e.g. the Examples 1.8 and 1.9 in [52]. An inspection of that proof shows that the occurring constants can be chosen independently of $\theta\in(0,1)$. We can now interpolate the results of Proposition 1.1 to derive

$$S_1\le c\big(M_{2+2\theta}^3+C(M_{2+2\theta},T)M_{2+2\theta}^2\big)\tau^{1+\theta}=:C_{1,1}\tau^{1+\theta} \tag{4.3}$$

with $C_{1,1}$ only depending on T and $M_{2+2\theta}$.

To deal with S_2 in (4.2), as in Section 3.2, we introduce the function

$$f:[0,T]\to H^2;\qquad f(t):=\big(B(u(t/2))-B(T(t/2)u_0)\big)T(t/2)u_0.$$

We write

$$\begin{aligned}f(t_1)-f(t_2)&=\big(B(u(t_1/2))-B(T(t_1/2)u_0)\big)\big(T(t_1/2)u_0-T(t_2/2)u_0\big)\\&\quad+\big(B(u(t_1/2))-B(u(t_2/2))\big)T(t_2/2)u_0\\&\quad-\big(B(T(t_1/2)u_0)-B(T(t_2/2)u_0)\big)T(t_2/2)u_0\end{aligned}$$

and estimate

$$\begin{aligned}\|f(t_1)-f(t_2)\|_{H^2}\le c\Big(&\big(\|u(t_1/2)\|_{H^2}^2+\|u_0\|_{H^2}^2\big)\,\|u_0\|_{H^{2+2\theta}}\\&+C(M_{2+2\theta},T)\big(\|u(t_1/2)\|_{H^2}+\|u(t_2/2)\|_{H^2}\big)\,\|u_0\|_{H^2}\\&+2\,\|u_0\|_{H^{2+2\theta}}\,\|u_0\|_{H^2}^2\Big)\cdot|t_1-t_2|^\theta\end{aligned}$$

for all $t_1,t_2\in[0,T]$, employing Lemma 2.1 and 2.3. Because of $f(0)=0$ we thus obtain

$$\begin{aligned}\|f(\tau)\|_{H^2}&\le c\big(M_2^2M_{2+2\theta}+C(M_{2+2\theta},T)M_2^2+M_{2+2\theta}M_2^2\big)\tau^\theta\\&\le c\big(M_{2+2\theta}^3+C(M_{2+2\theta},T)M_{2+2\theta}^2\big)\tau^\theta,\\S_2=\|\tau T(\tau/2)f(\tau)\|_{H^2}&\le C_{1,2}\tau^{1+\theta}\end{aligned} \tag{4.4}$$

with $C_{1,2}$ only depending on T and $M_{2+2\theta}$.

2) *Bound on I_2*: By means of Lemma 2.1, we bound the two summands of I_2 by

$$\begin{aligned}\Big\|\int_0^\tau T(\tau-s)B(u(s))\int_0^s T(s-\sigma)B(u(\sigma))u(\sigma)\,\mathrm{d}\sigma\,\mathrm{d}s\Big\|_{H^2}&\le c\tau^2M_2^5,\\\Big\|\int_0^\tau(\tau-s)T(\tau/2)B(u^*)^2e^{sB(u^*)}u^*\,\mathrm{d}s\Big\|_{H^2}&\le cM_2^3\tau^2.\end{aligned}$$

The assertion now follows by combining the above two inequalities with (4.1), (4.2), (4.3) and (4.4). □

4.2.2. Boundedness of the numerical solution in the H^2-norm

Proof (of Lemma 4.3):
Let $\theta \in (0,1)$. We denote by $u(s, y_0)$ the solution to (2.5) at time $s \geq 0$ with initial function $y_0 \in H^{2+2\theta}$. Let $u_0 \in H^{2+2\theta}$ and define

$$\tau_0 := \min\left\{ \left(\frac{M_2}{T e^{TC_2} C_1} \right)^{1/\theta}, T \right\}. \tag{4.5}$$

We prove with an induction argument part (b) and an even stronger version of part (a). For all $\tau \in (0, \tau_0]$, $n \in \mathbb{N}_0$ with $n\tau \leq T$ and $k \in \{0, \dots, n\}$ we claim

$$\left\| \Psi_\tau^{n-k}(u(k\tau, u_0)) - u(n\tau, u_0) \right\|_{H^2} \leq T e^{TC_2} C_1 \tau^\theta \tag{4.6}$$

with C_1 from Lemma 4.2 and C_2 from Lemma 3.4 (with $M := 2M_2$) and

$$\left\| \Psi_\tau^{n-k}(u(k\tau, u_0)) \right\|_{H^2} \leq 2M_2 =: \widehat{C}. \tag{4.7}$$

We first note that definition (4.5) and estimate (4.6) yield

$$\left\| \Psi_\tau^{n-k}(u(k\tau, u_0)) - u(n\tau, u_0) \right\|_{H^2} \leq M_2$$

for $\tau \in (0, \tau_0]$, so that the strong boundedness estimate (4.7) will follow from (4.6) via the triangle inequality.

We fix $\tau \in (0, \tau_0]$ and establish (4.6) by induction. The case $n = 0$ is trivial. Let the induction hypothesis

$$\left\| \Psi_\tau^{n-k}(u(k\tau, u_0)) - u(n\tau, u_0) \right\|_{H^2} \leq T e^{TC_2} C_1 \tau^\theta \leq M_2$$

hold true for all $k \in \{0, \dots, n\}$ and some $n \in \mathbb{N}_0$ with $(n+1)\tau \leq T$. Hence, (4.7) is valid for all $k \in \{0, \dots, n\}$. We now show (4.6) with n replaced by $n+1$. For $k = n+1$ the estimate (4.6) is clear. Let $k \in \{0, \dots, n\}$. Estimate (4.7) for n gives a uniform constant C_2 for following applications of Lemma 3.4, so that Lemma 3.4 and 4.2 imply with a telescopic sum that

$$\begin{aligned}
&\left\| \Psi_\tau^{n+1-k}(u(k\tau, u_0)) - u((n+1)\tau, u_0) \right\|_{H^2} \\
&\quad \leq \sum_{j=0}^{n-k} \left\| \Psi_\tau^{n-k-j}\big(\Psi_\tau\big(u((k+j)\tau, u_0)\big)\big) - \Psi_\tau^{n-k-j}\big(u((k+j+1)\tau, u_0)\big) \right\|_{H^2} \\
&\quad \leq \sum_{j=0}^{n-k} e^{(n-k-j)C_2\tau} \left\| \Psi_\tau\big(u((k+j)\tau, u_0)\big) - u\big(\tau, u((k+j)\tau, u_0)\big) \right\|_{H^2} \\
&\quad \leq \sum_{j=0}^{n-k} e^{(n-k-j)C_2\tau} C_1 \tau^{1+\theta} \leq \sum_{j=0}^{n-k} e^{C_2 T} C_1 \tau^{1+\theta} \leq T e^{TC_2} C_1 \tau^\theta.
\end{aligned}$$

Thereby, we can apply Lemma 4.2 with a uniform constant since we have

$$\sup_{t \in [0, T-l\tau]} \left\| u(t, u(l\tau, u_0)) \right\|_{H^{2+2\theta}} \leq M_{2+2\theta}$$

for all $l \in \{0, \dots, n\}$, compare Section 3.2. □

4.3. The estimate in L^2

We first prove Lemma 4.4. Afterwards we show Lemma 3.4 and combine it with Lemma 4.4 and 4.3 to infer Theorem 4.1.

4.3.1. The local error in the L^2-norm

The proof of Lemma 4.4 is similar to the one of Lemma 4.2, but we need a Taylor expansion of second order instead of first order. We use the following fact about a quadrature formula on a two-dimensional simplex.

Lemma 4.5. *Let $(X, \|\cdot\|)$ be a Banach space, $\tau > 0$ and*

$$S_\tau := \{(x, y) \in \mathbb{R}^2 \mid 0 \le y \le x \le \tau\}.$$

We define the linear operators

$$U_1 : C(S_\tau, X) \to X \quad \text{and} \quad U_2 : C^1(S_\tau, X) \to X$$

by

$$U_j f := \int_{S_\tau} f(x, y)\,\mathrm{d}(x, y) - \frac{\tau^2}{8}\big(f(0,0) + f(\tau,0) + f(\tau,\tau) + f(2\tau/3, \tau/3)\big).$$

These operators are bounded and we have

$$\|U_1 f\| \le \tau^2 \|f\|_C \quad \text{and} \quad \|U_2 f\| \le c\tau^3 \|f\|_{C^1}.$$

Proof:
The first estimate in the lemma is clear. To see the second one, we write

$$f(x, y) - f(a, b) = -\int_0^1 f'(x + r(a - x), y + r(b - y)) \cdot (a - x, b - y)\,\mathrm{d}r$$

for each $(a, b) \in \{(0,0), (\tau,0), (\tau,\tau), (2\tau/3, \tau/3)\}$. □

Proof (of Lemma 4.4):
Let $\theta \in (0, 1)$, $u_0 \in H^{2+2\theta}$ and $\tau \in (0, T]$. We first recall the representation (3.18) for $u(\tau) - \Psi_\tau(\tau)$,

$$\begin{aligned} u(\tau) &- \Psi_\tau(u_0) \\ &= \left(\int_0^\tau T(\tau - s)B(u(s))T(s)u_0\,\mathrm{d}s - \tau T(\tau/2)B(u^*)T(\tau/2)u_0\right) \\ &\quad + \left(\int_0^\tau T(\tau - s)B(u(s)) \int_0^s T(s - \sigma)B(u(\sigma))u(\sigma)\,\mathrm{d}\sigma\,\mathrm{d}s\right. \end{aligned}$$

$$
\begin{aligned}
&\quad - \frac{\tau^2}{2} T(\tau/2) B(u^*)^2 T(\tau/2) u_0 \Big) \\
&\quad - \frac{1}{2} \int_0^\tau (\tau - s)^2 T(\tau/2) B(u^*)^3 e^{sB(u^*)} u^* \, \mathrm{d}s \\
&=: I_1 + I_2 + I_3.
\end{aligned} \tag{4.8}
$$

1) *Bound on* I_1: We employ again the function $w : [0, T] \to L^2$ defined by

$$
w(s) := T(\tau - s) B(u(s)) T(s) u_0,
$$

see Section 3.3, and estimate (3.19), i.e.

$$
\begin{aligned}
\|I_1\|_{L^2} &\le \left\| \int_0^\tau T(\tau - s) B(u(s)) T(s) u_0 \, \mathrm{d}s - \tau w(\tau/2) \right\|_{L^2} \\
&\quad + \|\tau w(\tau/2) - \tau T(\tau/2) B(u^*) T(\tau/2) u_0\|_{L^2} \\
&=: S_1 + S_2.
\end{aligned} \tag{4.9}
$$

The first summand on the right-hand side will be controlled by interpolation. We already know from (3.5) that

$$
\begin{aligned}
w'(s) = &- T(\tau - s) AB(u(s)) T(s) u_0 \\
&- 2i\mu T(\tau - s) \operatorname{Re}\big(\overline{u(s)} Au(s)\big) T(s) u_0 \\
&+ T(\tau - s) B(u(s)) AT(s) u_0.
\end{aligned}
$$

This equality yields with Lemma 2.1 that

$$
\|w'(s)\|_{L^2} \le c\big(\|u(s)\|_{H^2}^2 \|u_0\|_{H^2} + \|u(s)\|_{H^2}^2 \|u_0\|_{H^2} + \|u(s)\|_{H^2}^2 \|u_0\|_{H^2}\big)
$$

and hence

$$
\sup_{s \in [0,T]} \|w'(s)\|_{L^2} \le c M_2^3.
$$

We have

$$
\begin{aligned}
w'(s_1) - w'(s_2) = &- T(\tau - s_1) AB(u(s_1)) \big(T(s_1) u_0 - T(s_2) u_0\big) \\
&- T(\tau - s_1) A\big(B(u(s_1)) - B(u(s_2))\big) T(s_2) u_0 \\
&- \big(T(\tau - s_1) - T(\tau - s_2)\big) AB(u(s_2)) T(s_2) u_0 \\
&- 2i\mu T(\tau - s_1) \operatorname{Re}\big(\overline{u(s_1)} Au(s_1)\big) \big(T(s_1) u_0 - T(s_2) u_0\big) \\
&- 2i\mu T(\tau - s_1) \operatorname{Re}\Big(\overline{u(s_1)} A\big(u(s_1) - u(s_2)\big)\Big) T(s_2) u_0 \\
&- 2i\mu T(\tau - s_1) \operatorname{Re}\Big(\big(\overline{u(s_1)} - \overline{u(s_2)}\big) Au(s_2)\Big) T(s_2) u_0 \\
&- 2i\mu \big(T(\tau - s_1) - T(\tau - s_2)\big) \operatorname{Re}\big(\overline{u(s_2)} Au(s_2)\big) T(s_2) u_0 \\
&+ T(\tau - s_1) B(u(s_1)) A\big(T(s_1) u_0 - T(s_2) u_0\big)
\end{aligned}
$$

$$
\begin{aligned}
&+ T(\tau - s_1)\big(B(u(s_1)) - B(u(s_2))\big)AT(s_2)u_0 \\
&+ \big(T(\tau - s_1) - T(\tau - s_2)\big)B(u(s_2))AT(s_2)u_0.
\end{aligned}
$$

Lemma 2.1 and 2.3 then imply

$$
\begin{aligned}
&\|w'(s_1) - w'(s_2)\|_{L^2} \\
&\quad\le c\Big(\|u(s_1)\|_{H^2}^2 \|u_0\|_{H^{2+2\theta}} + C(M_{2+2\theta}, T)\big(\|u(s_1)\|_{H^2} + \|u(s_2)\|_{H^2}\big)\|u_0\|_{H^2} \\
&\qquad + \|u(s_2)\|_{H^{2+2\theta}}^2 \|u_0\|_{H^{2+2\theta}} + \|u(s_1)\|_{H^2}^2 \|u_0\|_{H^{2+2\theta}} \\
&\qquad + \|u(s_1)\|_{H^2} C(M_{2+2\theta}, T)\|u_0\|_{H^2} + C(M_{2+2\theta}, T)\|u(s_1)\|_{H^2}\|u_0\|_{H^2} \\
&\qquad + \|u(s_2)\|_{H^2}\|u(s_2)\|_{H^{2+2\theta}}\|u_0\|_{H^2} + \|u(s_1)\|_{H^2}^2\|u_0\|_{H^{2+2\theta}} \\
&\qquad + \big(\|u(s_1)\|_{H^2} + \|u(s_2)\|_{H^2}\big)C(M_{2+2\theta}, T)\|u_0\|_{H^2} + \|u(s_2)\|_{H^2}^2\|u_0\|_{H^{2+2\theta}}\Big)\cdot \\
&\qquad \cdot |s_1 - s_2|^\theta
\end{aligned}
$$

for all $s_1, s_2 \in [0,T]$. The function w thus belongs to $C^{1,\theta}([0,T], L^2)$ and

$$
\begin{aligned}
\|w'(s_1) - w'(s_2)\|_{L^2} &\le c\big(M_2^2 M_{2+2\theta} + C(M_{2+2\theta}, T)M_2^2 + M_{2+2\theta}^3\big)|s_1 - s_2|^\theta \\
&\le c\big(M_{2+2\theta}^3 + C(M_{2+2\theta}, T)M_{2+2\theta}^2\big)|s_1 - s_2|^\theta
\end{aligned}
$$

for all $s_1, s_2 \in [0,T]$. Analogously as in the the proof of Lemma 4.2, $C^{1,\theta}([0,T], L^2)$ is the real interpolation space

$$
\big(C^1([0,T], L^2), C^2([0,T], L^2)\big)_{\theta,\infty}
$$

and the occurring constants are independent of $\theta \in (0,1)$. Hence, interpolation in Proposition 1.1 yields

$$
S_1 \le \tau^{2+\theta} c\Big(\big(M_{2+2\theta}^3 + C(M_{2+2\theta}, T)M_{2+2\theta}^2\big) + M_2^3\Big) =: C_{3,1}\tau^{2+\theta} \tag{4.10}
$$

with $C_{3,1}$ only depending on T and $M_{2+2\theta}$.

To treat the second summand in (4.2), we look at the function

$$
f : [0,T] \to L^2; \qquad f(t) := \big(B(u(t/2)) - B(T(t/2)u_0)\big)T(t/2)u_0,
$$

cf. (3.21). We want to check that f belongs to $C^{1,\theta}([0,T], L^2)$. Observe that

$$
\begin{aligned}
2f'(t) = &-2i\mu\Big(\mathrm{Re}\big(\overline{u}(t/2)Au(t/2)\big) - \mathrm{Re}\big(\overline{(T(t/2)u_0)}T(t/2)Au_0\big)\Big)T(t/2)u_0 \\
&+ \big(B(u(t/2)) - B(T(t/2)u_0)\big)T(t/2)Au_0.
\end{aligned}
$$

So, we have

$$
\begin{aligned}
2f'(t_1) - 2f'(t_2) = &-2i\mu\,\mathrm{Re}\big(\overline{u}(t_1/2)Au(t_1/2)\big)\big(T(t_1/2)u_0 - T(t_2/2)u_0\big) \\
&- 2i\mu\,\mathrm{Re}\big(\overline{(T(t_1/2)u_0)}T(t_1/2)Au_0\big)\big(T(t_1/2)u_0 - T(t_2/2)u_0\big)
\end{aligned}
$$

$$
\begin{aligned}
&- 2i\mu \operatorname{Re}\Big(\overline{u}(t_1/2)A\big(u(t_1/2)-u(t_2/2)\big)\Big)T(t_2/2)u_0\\
&- 2i\mu \operatorname{Re}\Big(\big(\overline{u}(t_1/2)-\overline{u}(t_2/2)\big)Au(t_2/2)\Big)T(t_2/2)u_0\\
&+ 2i\mu \operatorname{Re}\Big(\overline{(T(t_1/2)u_0)}A\big(T(t_1/2)u_0-T(t_2/2)u_0\big)\Big)T(t_2/2)u_0\\
&+ 2i\mu \operatorname{Re}\Big(\big(\overline{(T(t_1/2)u_0)}-\overline{(T(t_2/2)u_0)}\big)T(t_2/2)Au_0\Big)T(t_2/2)u_0\\
&+ \big(B(u(t_1/2))-B(T(t_1/2)u_0)\big)A\big(T(t_1/2)u_0-T(t_2/2)u_0\big)\\
&+ \big(B(u(t_1/2))-B(u(t_2/2))\big)T(t_2/2)Au_0\\
&- \big(B(T(t_1/2)u_0)-B(T(t_2/2)u_0)\big)T(t_2/2)u_0
\end{aligned}
$$

and thus

$$
\begin{aligned}
&\|f'(t_1)-f'(t_2)\|_{L^2}\\
&\quad\le c\Big(\|u(t_1/2)\|_{H^2}^2\,\|u_0\|_{H^{2+2\theta}}+\|u_0\|_{H^2}^2\,\|u_0\|_{H^{2+2\theta}}\\
&\qquad+\|u(t_1/2)\|_{H^2}\,C(M_{2+2\theta},T)\,\|u_0\|_{H^2}+C(M_{2+2\theta},T)\,\|u(t_2/2)\|_{H^2}\,\|u_0\|_{H^2}\\
&\qquad+\|u_0\|_{H^2}\,\|u_0\|_{H^{2+2\theta}}\,\|u_0\|_{H^2}+\|u_0\|_{H^{2+2\theta}}\,\|u_0\|_{H^2}^2\\
&\qquad+\big(\|u(t_1/2)\|_{H^2}^2+\|u_0\|_{H^2}^2\big)\,\|u_0\|_{H^{2+2\theta}}\\
&\qquad+\big(\|u(t_1/2)\|_{H^2}+\|u(t_2/2)\|_{H^2}\big)C(M_{2+2\theta},T)\,\|u_0\|_{H^2}\\
&\qquad+2\,\|u_0\|_{H^2}\,\|u_0\|_{H^{2+2\theta}}\,\|u_0\|_{H^2}\Big)\cdot|t_1-t_2|^\theta\\
&\le c\big(M_2^2M_{2+2\theta}+M_2^2C(M_{2+2\theta},T)\big)\,|t_1-t_2|^\theta\\
&\le c\big(M_{2+2\theta}^3+M_{2+2\theta}^2C(M_{2+2\theta},T)\big)\,|t_1-t_2|^\theta\\
&=: C_{3,2}\,|t_1-t_2|^\theta
\end{aligned}
$$

for all $t_1,t_2\in[0,T]$, with a constant $C_{3,2}$ only depending on T and $M_{2+2\theta}$. Together with $f(0)=0$ and $f'(0)=0$ this inequality implies

$$
\begin{aligned}
\|f(\tau)\|_{L^2}\le\Big\|\int_0^\tau\big(f'(s)-f'(0)\big)\,\mathrm{d}s\Big\|_{L^2}&\le C_{3,2}\tau^{1+\theta},\\
S_2=\|\tau T(\tau/2)f(\tau)\|_{L^2}&\le C_{3,2}\tau^{2+\theta}.
\end{aligned}
\tag{4.11}
$$

2) *Bound on I_2*: We now tackle the summand I_2 in (4.8). We define, as in (3.24), the function $v:[0,T]\times[0,T]\to L^2$ by

$$
v(s,\sigma):=T(\tau-s)B(u(s))T(s-\sigma)B(u(\sigma))u(\sigma)
$$

and split

$$
\|I_2\|_{L^2}\le\Big\|\int_0^\tau\int_0^s T(\tau-s)B(u(s))T(s-\sigma)B(u(\sigma))u(\sigma)\,\mathrm{d}\sigma\,\mathrm{d}s
$$

$$
\begin{aligned}
&\qquad - \frac{\tau^2}{8}\big(v(0,0)+v(\tau,0)+v(\tau,\tau)+v(2\tau/3,\tau/3)\big)\Big\|_{L^2},\\
&+ \quad \Big\|\frac{\tau^2}{8}\big(v(0,0)+v(\tau,0)+v(\tau,\tau)+v(2\tau/3,\tau/3)\big)\\
&\qquad - \frac{\tau^2}{2}T(\tau/2)B(u^*)^2T(\tau/2)u_0\Big\|_{L^2}\\
&=: R_1+R_2,
\end{aligned}
\tag{4.12}
$$

as in (3.25). For all $(s_1,\sigma_1),(s_2,\sigma_2)\in S_\tau$ we have

$$
\begin{aligned}
&v(s_1,\sigma_1)-v(s_2,\sigma_2)\\
&\quad = T(\tau-s_1)B(u(s_2))T(s_1-\sigma_1)B(u(\sigma_1))\big(u(\sigma_1)-u(\sigma_2)\big)\\
&\qquad - T(\tau-s_1)B(u(s_1))T(s_1-\sigma_1)\big(B(u(\sigma_1))-B(u(\sigma_2))\big)u(\sigma_2)\\
&\qquad + T(\tau-s_1)B(u(s_2))\big(T(s_1-\sigma_1)-T(s_2-\sigma_2)\big)B(u(\sigma_2))u(\sigma_2)\\
&\qquad + T(\tau-s_1)\big(B(u(s_1))-B(u(s_2))\big)T(s_2-\sigma_2)B(u(\sigma_2))u(\sigma_2)\\
&\qquad + \big(T(\tau-s_1)-T(\tau-s_2)\big)B(u(s_2))T(s_2-\sigma_2)B(u(\sigma_2))u(\sigma_2).
\end{aligned}
$$

Lemma 2.1 and 2.3 then yield

$$
\begin{aligned}
&\|v(s_1,\sigma_1)-v(s_2,\sigma_2)\|_{L^2}\\
&\le c\Big(\|u(s_1)\|_{H^2}^2\,\|u(\sigma_1)\|_{H^2}^2\,C(M_{2+2\theta},T)\,|\sigma_1-\sigma_2|^\theta\\
&\quad + \|u(s_1)\|_{H^2}^2\,\big(\|u(\sigma_1)\|_{H^2}+\|u(\sigma_2)\|_{H^2}\big)C(M_{2+2\theta},T)\,\|u(\sigma_2)\|_{H^2}\,|\sigma_1-\sigma_2|^\theta\\
&\quad + \|u(s_2)\|_{H^2}^2\,\|u(\sigma_2)\|_{H^{2+2\theta}}^3\,|s_1-s_2+\sigma_1-\sigma_2|^\theta\\
&\quad + \big(\|u(s_1\|_{H^2}+\|u(s_2)\|_{H^2}\big)C(M_{2+2\theta},T)\,\|u(\sigma_2)\|_{H^2}^3\,|s_1-s_2|^\theta\\
&\quad + \|u(s_2)\|_{H^2}^2\,\|u(\sigma_2)\|_{H^2}^2\,\|u(\sigma_2)\|_{H^{2\theta}}\,|s_1-s_2|^\theta\Big),\\
&\le c\big(M_2^4C(M_{2+2\theta},T)+M_2^2M_{2+2\theta}^3+M_2^4M_{2\theta}\big)\left|\begin{pmatrix} s_1-s_2\\ \sigma_1-\sigma_2\end{pmatrix}\right|^\theta\\
&\le c\big(M_{2+2\theta}^4C(M_{2+2\theta},T)+M_{2+2\theta}^5\big)\left|\begin{pmatrix} s_1-s_2\\ \sigma_1-\sigma_2\end{pmatrix}\right|^\theta\\
&\le C_{4,1}\left|\begin{pmatrix} s_1-s_2\\ \sigma_1-\sigma_2\end{pmatrix}\right|^\theta
\end{aligned}
$$

with $C_{4,1}$ only depending on T and $M_{2+2\theta}$. By interpolating in Lemma 4.5, we obtain as above the inequality

$$
R_1\le c\tau^{2+\theta}C_{4,1}. \tag{4.13}
$$

To estimate R_2, we introduce the function $g:[0,T]\to L^2$ by

$$
g(t):=T(t)B(u_0)^2u_0+B(u(t))T(t)B(u_0)u_0
$$

$$
\begin{aligned}
&+ B(u(t))^2 u(t) + T(t/3)B(u(2t/3))T(t/3)B(u(t/3))u(t/3) \\
&- 4T(t/2)B(T(t/2)u_0)^2 T(t/2)u_0.
\end{aligned}
$$

For all $t_1, t_2 \in [0,T]$ we have

$$
\begin{aligned}
&g(t_1) - g(t_2) \\
&\quad = \big(T(t_1) - T(t_2)\big)B(u_0)^2 u_0 \\
&\qquad + B(u(t_1))\big(T(t_1)B(u_0)u_0 - T(t_2)B(u_0)u_0\big) \\
&\qquad + \big(B(u(t_1)) - B(u(t_2))\big)T(t_2)B(u_0)u_0 \\
&\qquad + B(u(t_1))^2\big(u(t_1) - u(t_2)\big) \\
&\qquad + \big(B(u(t_1)) + B(u(t_2))\big)\big(B(u(t_1)) - B(u(t_2))\big)u(t_2) \\
&\qquad + T(t_1/3)B(u(2t_1/3))T(t_1/3)B(u(t_1/3))\big(u(t_1/3) - u(t_2/3)\big) \\
&\qquad + T(t_1/3)B(u(2t_1/3))T(t_1/3)\big(B(u(t_1/3)) - B(u(t_2/3))\big)u(t_2/3) \\
&\qquad + T(t_1/3)B(u(2t_1/3))\big(T(t_1/3) - T(t_2/3)\big)[B(u(t_2/3))u(t_2/3)] \\
&\qquad + T(t_1/3)\big(B(u(2t_1/3)) - B(u(2t_2/3))\big)T(t_2/3)B(u(t_2/3))u(t_2/3) \\
&\qquad + \big(T(t_1/3) - T(t_2/3)\big)B(u(2t_2/3))T(t_2/3)B(u(t_2/3))u(t_2/3) \\
&\qquad - 4T(t_1/2)B(T(t_1/2)u_0)^2\big(T(t_1/2)u_0 - T(t_2/2)u_0\big) \\
&\qquad + T(t_1/2)\big(B(T(t_1/3)u_0) + B(T(t_2/3)u_0)\big) \\
&\qquad\quad \big(B(T(t_1/3)u_0) - B(u(t_2/3)u_0)\big)T(t_2/2)u_0 \\
&\qquad - 4\big(T(t_1/2) - T(t_2/2)\big)B(T(t_2/2)u_0)^2 T(t_2/2)u_0.
\end{aligned}
$$

From Lemma 2.1 and 2.3 we thus derive

$$
\begin{aligned}
&\|g(t_1) - g(t_2)\|_{L^2} \\
&\quad \le c\Big(\|u_0\|_{H^2}^4 \|u_0\|_{H^{2\theta}} + \|u(t_1)\|_{H^2}^2 \|u_0\|_{H^2}^2 \|u_0\|_{H^{2\theta}} \\
&\qquad + \big(\|u(t_1)\|_{L^2} + \|u(t_2)\|_{L^2}\big)C(M_{2+2\theta},T)\,\|u_0\|_{H^2}^3 \\
&\qquad + \|u(t_1)\|_{H^2}^3 \|u(t_1)\|_{L^2} C(M_{2+2\theta},T) \\
&\qquad + \big(\|u(t_1)\|_{H^2}^2 + \|u(t_2)\|_{H^2}^2\big)\big(\|u(t_1)\|_{H^2} + \|u(t_2)\|_{H^2}\big)C(M_{2+2\theta},T)\,\|u(t_2)\|_{L^2} \\
&\qquad + \|u(2t_1/3)\|_{H^2}^2 \|u(t_1/3)\|_{H^2} \|u(t_1/3)\|_{L^2} C(M_{2+2\theta},T) \\
&\qquad + \|u(2t_1/3)\|_{H^2}^2 \big(\|u(t_1/3)\|_{H^2} + \|u(t_2/3)\|_{H^2}\big)C(M_{2+2\theta},T)\,\|u(t_2/3)\|_{L^2} \\
&\qquad + \|u(2t_1/3)\|_{H^2}^2 \|u(t_2/3)\|_{H^2}^2 \|u(t_2/3)\|_{H^{2\theta}} \\
&\qquad + \big(\|u(2t_1/3)\|_{H^2} + \|u(2t_2/3)\|_{H^2}\big)C(M_{2+2\theta},T)\,\|u(t_2/3)\|_{H^2}^2 \|u(t_2/3)\|_{L^2} \\
&\qquad + \|u(2t_2/3)\|_{H^2}^2 \|u(t_2/3)\|_{H^2}^2 \|u(t_2/3)\|_{H^{2\theta}} \\
&\qquad + \|u_0\|_{H^2}^4 \|u_0\|_{H^{2\theta}} \\
&\qquad + 2\,\|u_0\|_{H^2}^2 \big(\|u(t_1/2)\|_{H^2} + \|u(t_2/2)\|_{H^2}\big)C(M_{2+2\theta},T)\,\|u_0\|_{L^2} \\
&\qquad + \|u_0\|_{H^2}^4 \|u_0\|_{H^{2\theta}}\Big) \cdot |t_1 - t_2|^\theta
\end{aligned}
$$

$$\begin{aligned}
&\leq c\big(M_2^3 M_0 C(M_{2+2\theta}, T) + M_2^4 M_{2\theta}\big)\,|t_1 - t_2|^\theta \\
&\leq c\big(M_2^5 + M_2^4 C(M_{2+2\theta}, T)\big)\,|t_1 - t_2|^\theta \\
&\leq C_{4,2}\,|t_1 - t_2|^\theta
\end{aligned}$$

with $C_{4,2}$ only depending on T and $M_{2+2\theta}$. Because of $g(0) = 0$, this inequality leads to the bound

$$R_2 = \left\| \frac{\tau^2}{8} g(\tau) \right\|_{L^2} \leq C_{4,2} \tau^{2+\theta}. \tag{4.14}$$

3) *Bound on* I_3: Lemma 2.1 implies that

$$\left\| \frac{1}{2} \int_0^\tau (\tau - s)^2 T(\tau/2) B(u^*)^3 e^{sB(u^*)} u^* \,\mathrm{d}s \right\|_{L^2} \leq c M_2^7 \tau^3.$$

This estimate and (4.8), (4.9), (4.10), (4.11), (4.12), (4.13) and (4.14) imply the assertion. □

4.3.2. Convergence in the L^2-norm

Proof (of Theorem 4.1):
Let $\theta > 0$ and $u_0 \in H^{2+2\theta}$. Take $\tau \in (0, \tau_0]$ with the bound $\tau_0 > 0$ on the time step size from Lemma 4.3 and $n \in \mathbb{N}$ with $n\tau \leq T$. We have

$$u(n\tau) - \Psi_\tau^n(u_0) = \sum_{k=0}^{n-1} \Big(\Psi_\tau^k\big(u((n-k)\tau)\big) - \Psi_\tau^{k+1}\big(u((n-k-1)\tau)\big) \Big).$$

In view of Lemma 4.3, the expressions $\Psi_\tau^l\big(u((n-1-l)\tau)\big)$ with $l \in \{0, \dots, n-1\}$ are bounded in H^2 by a constant $\widehat{C}$ that only depends on M_2 (and not on n or τ). Iteratively, Lemma 3.8 can thus be applied with $M := \widehat{C}$ to all summands appearing in the second line of the following calculation. Together with Lemma 4.4 we derive

$$\begin{aligned}
&\|u(n\tau) - \Psi_\tau^n(u_0)\|_{L^2} \\
&\quad\leq \sum_{k=0}^{n-1} \big\| \Psi_\tau^k\big(u((n-k)\tau)\big) - \Psi_\tau^{k+1}\big(u((n-k-1)\tau)\big) \big\|_{L^2} \\
&\quad\leq \sum_{k=0}^{n-1} e^{kC_4\tau} \big\| u\big(\tau, u((n-k-1)\tau)\big) - \Psi_\tau\big(u((n-k-1)\tau)\big) \big\|_{L^2} \\
&\quad\leq \sum_{k=0}^{n-1} e^{kC_4\tau} C_3 \tau^{2+\theta} \leq \sum_{k=0}^{n_1} e^{C_4 T} C_3 \tau^{2+\theta} \leq T e^{TC_4} C_3 \tau^{1+\theta}.
\end{aligned}$$

As in the analogous situation in the proof of Lemma 4.3 we can apply Lemma 4.4 with a uniform constant C_3 since

$$\sup_{t \in [0, T - l\tau]} \|u(t, u(l\tau, u_0))\|_{H^2} \leq M_4.$$

This completes the proof of Theorem 4.1. □

5. Convergence of the Strang and the Lie splitting for initial functions in H^2

In this chapter we extend our analysis from Chapter 4 to the situation that the initial function is only in H^2. In contrast to in Chapter 3 and 4 we investigate not only the Strang splitting but also the Lie splitting. In Section 5.1 we state that they both converge with order one.

The main problem in transferring the proof from Chapter 4 to this situation with initial functions of low regularity is that the order of the local error in the H^2-norm is no longer strictly larger than (but equal to) one. This implies that the proofs of the analoga to Lemma 4.3 on the strong boundedness, see Lemma 5.7 and 5.8, cannot be carried out as before. Moreover, these lemmas cannot be omitted completely since the error constant in the final part of the proof is not allowed to depend on the time step size and the time step. The remedy is to use interpolation in the domains to show the strong boundedness. We give the details of this procedure in Section 5.2.

It is a natural question to ask if initial functions in other H^s-spaces are also worth to look at. One can lower the regularity below H^2 with the drawback that the solution to (2.5), or at least the derivative of this solution, is in a distributional space H^{-r} for some $r > 0$. We did not investigate that situation in this thesis. Theorem 3.1 and 5.2 show that the Strang and the Lie splitting have their classical order for initial functions in H^4 and H^2, respectively. Since this order cannot be improved, the investigation with initial functions with higher regularity does not lead to new interesting result, except one measures the errors in an H^s-norm for an $s > 0$.

5.1. The theorems for initial functions in H^2

The main results of this chapter are the following two convergence theorems for the Strang and the Lie splitting.

Theorem 5.1. *For each $u_0 \in H^2$ there exists a bound $\tau_0 > 0$ on the time step size such*

that we have

$$\|u(n\tau) - \Psi_\tau^n(u_0)\|_{L^2} \leq C\tau$$

for all $\tau \in (0, \tau_0]$ *and* $n \in \mathbb{N}$ *with* $n\tau \leq T$ *with a constant* $C \geq 0$ *that depends only on* u_0 *and* T*. More precisely,* C *depends only on* T *and* M_2*.*

The number $\tau_0 = \tau_0(T, M_2)$ is given in Lemma 5.7.

Theorem 5.2. *For each* $u_0 \in H^2$ *there exists a bound* $\tau_0 > 0$ *on the time step size such that we have*

$$\|u(n\tau) - \Phi_\tau^n(u_0)\|_{L^2} \leq C\tau$$

for all $\tau \in (0, \tau_0]$ *and* $n \in \mathbb{N}$ *with* $n\tau \leq T$ *with a constant* $C \geq 0$ *that depends only on* u_0 *and* T*. More precisely,* C *depends only on* T *and* M_2*.*

The number $\tau_0 = \tau_0(T, M_2)$ is given in Lemma 5.8.

The proofs of Theorem 5.1 and 5.2 are similar to the one of Theorem 4.1. The main difference is that in the first part of the proofs the local errors are estimated not in the H^2-norm but in the $H^{7/4}$-norm, see Lemma 5.3 and 5.4. This has the advantage that we obtain a local error of order 9/8 instead of order one. Additionally, $H^{7/4}$ is still an algebra due to Lemma 2.1. So, the stability estimates in $H^{7/4}$, see Lemma 5.5 and 5.6, can be shown in the same way as the stability estimate in Lemma 3.4. Since $9/8 > 1$, we can then prove Lemma 5.7 and 5.8 in the same way as Lemma 4.3 (with $\theta = 1/8$).

The proofs of the next two lemmas concerning the local errors for the Strang and the Lie splitting are discussed and shown in Section 5.2.

Lemma 5.3. *For all* $u_0 \in H^2$ *and* $\tau \in (0, T]$ *we have*

$$\begin{aligned} \|u(\tau) - \Psi_\tau(u_0)\|_{H^{7/4}} &\leq C_1\tau^{9/8}, \\ \|u(\tau) - \Psi_\tau(u_0)\|_{L^2} &\leq C_3\tau^2, \end{aligned}$$

with constants $C_1, C_3 \geq 0$ *depending only on* T *and* M_2*.*

Lemma 5.4. *For all* $u_0 \in H^2$ *and* $\tau \in (0, T]$ *we have*

$$\begin{aligned} \|u(\tau) - \Phi_\tau(u_0)\|_{H^{7/4}} &\leq C_5\tau^{9/8}, \\ \|u(\tau) - \Phi_\tau(u_0)\|_{L^2} &\leq C_7\tau^2, \end{aligned}$$

with constants $C_5, C_7 \geq 0$ *depending only on* T *and* M_2*.*

As explained above, the proof of the following stability, convergence and boundedness properties can be seen in the same way as in Chapter 3 and 4.

Lemma 5.5. *Let $M \geq 0$ and $u_0, v_0 \in H^2$ with $\|u_0\|_{H^{7/4}} \leq M$ and $\|v_0\|_{H^{7/4}} \leq M$. Then there are constants $C_2, C_4 \geq 0$, only depending on M, such that*

$$\begin{aligned} \|\Psi_\tau(u_0) - \Psi_\tau(v_0)\|_{H^{7/4}} &\leq e^{C_2\tau} \|u_0 - v_0\|_{H^{7/4}}, \\ \|\Psi_\tau(u_0) - \Psi_\tau(v_0)\|_{L^2} &\leq e^{C_4\tau} \|u_0 - v_0\|_{L^2} \end{aligned}$$

for all $\tau \in (0, T]$.

Lemma 5.6. *Let $M \geq 0$ and $u_0, v_0 \in H^2$ with $\|u_0\|_{H^{7/4}} \leq M$ and $\|v_0\|_{H^{7/4}} \leq M$. Then there are constants $C_6, C_8 \geq 0$, only depending on M, such that*

$$\begin{aligned} \|\Phi_\tau(u_0) - \Phi_\tau(v_0)\|_{H^{7/4}} &\leq e^{C_6\tau} \|u_0 - v_0\|_{H^{7/4}}, \\ \|\Phi_\tau(u_0) - \Phi_\tau(v_0)\|_{L^2} &\leq e^{C_8\tau} \|u_0 - v_0\|_{L^2} \end{aligned}$$

for all $\tau \in (0, T]$.

Lemma 5.7. *Let $u_0 \in H^2$. There exists a bound $\tau_0 > 0$ on the time step size, which is given by*

$$\tau_0 := \min\left\{ \left(\frac{M_{7/4}}{T e^{TC_6} C_5} \right)^8, T \right\},$$

with C_5 from Lemma 5.3 and C_6 from Lemma 5.5, such that the following two statements hold true.

(a) For all $\tau \in (0, \tau_0]$ and $n \in \mathbb{N}$ with $n\tau \leq T$ we have

$$\|\Psi_\tau^n(u_0) - u(n\tau)\|_{H^{7/4}} \leq C\tau^{1/8},$$

with a constant $C \geq 0$ depending only on T and M_2, i.e. the Strang splitting converges in $H^{7/4}$ with order $1/8$.

(b) Ψ_τ is strongly bounded for (2.5) in $H^{7/4}$ for initial functions in H^2, i.e. there exists a constant $\widehat{C} \geq 0$, only depending on M_2, such that $\left\|\Psi_\tau^{n-k}(u(k\tau))\right\|_{H^{7/4}} \leq \widehat{C}$ for all $\tau \in (0, \tau_0]$ and $n \in \mathbb{N}$ with $n\tau \leq T$ and $k \in \{0, \ldots, n\}$. In particular, the numerical solution is bounded in $H^{7/4}$ (choose $k = 0$).

Lemma 5.8. *Let $u_0 \in H^2$. There exists a bound $\tau_0 > 0$ on the time step size, which is given by*

$$\tau_0 := \min\left\{ \left(\frac{M_{7/4}}{T e^{TC_6} C_5} \right)^8, T \right\},$$

with C_5 from Lemma 5.4 and C_6 from Lemma 5.6, such that the following two statements hold true.

5. Convergence of the Strang and the Lie splitting for initial functions in H^2

(a) For all $\tau \in (0, \tau_0]$ and $n \in \mathbb{N}$ with $n\tau \leq T$ we have

$$\|\Phi_\tau^n(u_0) - u(n\tau)\|_{H^{7/4}} \leq C\tau^{1/8},$$

with a constant $C \geq 0$ depending only on T and M_2, i.e. the Lie splitting converges in $H^{7/4}$ with order $1/8$.

(b) Φ_τ is strongly bounded for (2.5) in $H^{7/4}$ for initial functions in H^2, i.e. there exists a constant $\widehat{C} \geq 0$, only depending on M_2, such that $\left\|\Phi_\tau^{n-k}(u(k\tau))\right\|_{H^{7/4}} \leq \widehat{C}$ for all $\tau \in (0, \tau_0]$ and $n \in \mathbb{N}$ with $n\tau \leq T$ and $k \in \{0, \ldots, n\}$. In particular, the numerical solution is bounded in $H^{7/4}$ (choose $k = 0$).

Remark 5.9. *Theorem 5.1 can be seen as an extension of Theorem 4.1 to the case $\theta = 0$. Of course, this fact is not interesting for applications since the simpler Lie splitting also converges with order one in L^2 due to Theorem 5.2. However, one can use Lemma 5.3 and 5.5 for an alternative proof of Theorem 4.1 but with the bound τ_0 from Lemma 5.7, which does not depend on θ. We omit the details of the proofs of these claims.*

5.2. The proofs of the theorems

We first prove Lemma 5.4. The proof of Lemma 5.3 can be done similarly. We start with an interpolation lemma that is closely related to Proposition 1.1. The very simple proof is omitted.

Lemma 5.10. *Let $T > 0$ and $\tau \in (0, T]$. We define the Banach space*

$$Z := C^1([0,T], L^2) \cap C([0,T], H^2)$$

with norm

$$\|f\|_Z := \|f\|_{C^1([0,T],L^2)} + \|f\|_{C([0,T],H^2)}$$

and the linear operators

$$V_1 : Z \to H^2 \quad \text{and} \quad V_2 : Z \to L^2 \quad \text{by} \quad V_j f := \int_0^\tau f(s)\,\mathrm{d}s - \tau f(0).$$

These operators are bounded and we have

$$\|V_1 f\|_{H^2} \leq 2\tau \|f\|_Z \qquad \text{and} \qquad \|V_2 f\|_{L^2} \leq \tau^2 \|f\|_Z .$$

Proof (of Lemma 5.4):
Let $u_0 \in H^2$ and $\tau > 0$. By Theorem 1.9, the solution to (2.5) at time τ is given by

$$u(\tau) = T(\tau)u_0 + \int_0^\tau T(\tau - s)B(u(s))u(s)\,\mathrm{d}s, \tag{5.1}$$

see (2.11). Applying the Taylor expansion

$$e^{\tau x} = 1 + \tau x + \int_0^\tau (\tau - s) x^2 e^{\tau x} \, \mathrm{d}s$$

to $\Phi_\tau(u_0) = \exp\big(\tau B(\widetilde{u})\big)\widetilde{u}$ with $\widetilde{u} = T(\tau)u_0$, see definition (2.9), we determine the numerical solution after one time step as

$$\Phi_\tau(u_0) = T(\tau)u_0 + \tau B(\widetilde{u})\widetilde{u} + \int_0^\tau (\tau - s) B(\widetilde{u})^2 e^{sB(\widetilde{u})}\widetilde{u} \, \mathrm{d}s. \tag{5.2}$$

The difference of (5.1) and (5.2) is

$$\begin{aligned} u(\tau) - \Phi_\tau(u_0) &= \left(\int_0^\tau T(\tau - s) B(u(s)) u(s) \, \mathrm{d}s - \tau B(\widetilde{u})\widetilde{u} \right) \\ &\quad - \int_0^\tau (\tau - s) B(\widetilde{u})^2 e^{sB(\widetilde{u})}\widetilde{u} \, \mathrm{d}s \\ &=: I_1 + I_2. \end{aligned} \tag{5.3}$$

1) *Bound on* I_1: We again look at the function

$$w : [0, T] \to H^2; \qquad w(s) := T(\tau - s) B(u(s)) u(s).$$

We abbreviate

$$S_1 := \int_0^\tau T(\tau - s) B(u(s)) u(s) \, \mathrm{d}s - \tau w(0) \qquad \text{and} \qquad S_2 := \tau w(0) - \tau B(\widetilde{u})\widetilde{u},$$

and write I_1 as the telescopic sum

$$\begin{aligned} I_1 &= \left(\int_0^\tau T(\tau - s) B(u(s)) u(s) \, \mathrm{d}s - \tau w(0) \right) + \Big(\tau w(0) - \tau B(\widetilde{u})\widetilde{u} \Big) \\ &= S_1 + S_2. \end{aligned} \tag{5.4}$$

With identity (3.4) and problem (2.5) we see that the derivative of w is

$$\begin{aligned} w'(s) = &- T(\tau - s) A B(u(s)) u(s) \\ &- 2i\mu T(\tau - s) \operatorname{Re}\big(\overline{u}(s) A u(s)\big) u(s) \\ &+ T(\tau - s) B(u(s)) \big(A u(s) + B(u(s)) u(s) \big). \end{aligned}$$

Lemma 2.1 now implies

$$\begin{aligned} \|w(s)\|_{H^2} &\le c \, \|u(s)\|_{H^2}^3 , \\ \|w(s)\|_{L^2} &\le c \, \|u(s)\|_{H^2}^2 \, \|u(s)\|_{L^2} , \\ \|w'(s)\|_{L^2} &\le c \big(\|u(s)\|_{H^2}^3 + \|u(s)\|_{H^2}^3 + \|u(s)\|_{H^2}^3 + \|u(s)\|_{H^2}^4 \, \|u(s)\|_{L^2} \big) \end{aligned}$$

for all $s \in [0,T]$. We thus obtain

$$\begin{aligned}
&\sup_{s\in[0,T]} \|w(s)\|_{H^2} \le cM_2^3,\\
&\sup_{s\in[0,T]} \|w(s)\|_{L^2} \le c\big(M_2^2 M_0\big) \le cM_2^3,\\
&\sup_{s\in[0,T]} \|w'(s)\|_{L^2} \le c\big(M_2^3 + M_2^4 M_0\big) \le c\big(M_2^3 + M_2^5\big).
\end{aligned}$$

By these inequalities, w belongs to $C^1([0,T],L^2)\cap C([0,T],H^2)$ and its norm in this space is bounded by a constant $C_{1,1}$ only depending on M_2. Lemma 5.10 then gives

$$\|S_1\|_{L^2} \le C_{1,1}\tau^2 \qquad \text{and} \qquad \|S_1\|_{H^2} \le 2C_{1,1}\tau. \tag{5.5}$$

Additionally, by interpolation w is contained in $C^{0,1/8}([0,T],H^{7/4})$ and

$$\|S_1\|_{H^{7/4}} \le cC_{1,1}\tau^{9/8}. \tag{5.6}$$

For the estimation of S_2 we first note that

$$S_2 = \tau T(\tau)B(u_0)u_0 - \tau B(T(\tau)u_0)T(\tau)u_0.$$

We define the function $f : [0,T] \to H^2$ by

$$f(t) := T(t)B(u_0)u_0 - B(T(t)u_0)T(t)u_0.$$

Since

$$\begin{aligned}
f(t_1) - f(t_2) = {}& \big(T(t_1)B(u_0)u_0 - T(t_2)B(u_0)u_0\big)\\
&- B(T(t_1)u_0)\big(T(t_1)u_0 - T(t_2)u_0\big)\\
&- \big(B(T(t_1)u_0) - B(T(t_2)u_0)\big)\big)T(t_2)u_0,
\end{aligned}$$

we deduce (with $\theta = 1/8$) from Lemma 2.1 and 2.3

$$\begin{aligned}
\|f(t_1) - f(t_2)\|_{H^{7/4}} &\le c\big(\|u_0\|_{H^2}^3 + \|u_0\|_{H^{7/4}}^2 \|u_0\|_{H^2}\\
&\qquad + 2\,\|u_0\|_{H^{7/4}} \|u_0\|_{H^2} \|u_0\|_{H^{7/4}}\big) \cdot |t_1 - t_2|^{1/8}\\
&\le c\big(M_2^3 + M_{7/4}^2 M_2\big)\,|t_1 - t_2|^{1/8}\\
&\le cM_2^3\,|t_1 - t_2|^{1/8}
\end{aligned}$$

for all $t_1, t_2 \in [0,T]$. Due to $f(0) = 0$, we thus have

$$\|S_2\|_{H^{7/4}} = \|\tau f(\tau)\|_{H^{7/4}} \le cM_2^3\tau^{9/8} \le C_{1,2,74}\tau^{9/8} \tag{5.7}$$

with a constant $C_{1,2,74}$ only depending on M_2. The derivative of f is given by

$$f'(t) = T(t)AB(u_0)u_0 + 2i\mu\,\mathrm{Re}\big(\overline{(T(t)u_0)}AT(t)u_0\big)T(t)u_0 - B(T(t)u_0)AT(t)u_0.$$

As before we can estimate this by

$$\|f'(t)\|_{L^2} \leq c\big(\|u_0\|_{H^2}^3 + \|u_0\|_{L^2}\,\|u_0\|_{H^2}^2 + \|u_0\|_{H^2}^3\big)$$

for all $t \in [0,T]$, which yields

$$\sup_{t\in[0,T]} \|f'(t)\|_{L^2} \leq c\big(M_2^3 + M_2^2 M_0\big) \leq cM_2^3.$$

Again due to $f(0) = 0$, it follows

$$f(\tau) = \int_0^\tau f'(s)\,\mathrm{d}s$$

and thus with the above estimate

$$\|S_2\|_{L^2} = \|\tau f(\tau)\|_{L^2} \leq C_{1,2,0}\tau^2 \tag{5.8}$$

with a constant $C_{1,2,0}$ only depending on M_2.

2) *Bound on* I_2: Lemma 2.1 allows us to bound the term I_2 in (5.3) by

$$\|I_2\|_{H^{7/4}} \leq cM_2^5 T^{7/8}\tau^{9/8} \qquad \text{and} \qquad \|I_2\|_{L^2} \leq cM_2^5\tau^2.$$

The proof now is finished by combing the estimates of I_2 with (5.3), (5.4), (5.6), (5.5), (5.7) and (5.8). □

Proof (of Theorem 5.1 and 5.2):
The stability properties of Lie splitting, see Lemma 5.6, are shown in the same manner as the ones for the Strang splitting in Lemma 3.4 and 3.8, but with H^2 replaced by $H^{7/4}$. Analogously as in Lemma 3.6 we see the strong boundedness of the numerical solution in Lemma 5.8. Then we deduce Theorem 5.2 by combining Lemma 5.4, 5.6 and 5.8 with the same technique as in the proof of Theorem 3.1. Theorem 5.1 is shown in the same way.□

6. Numerical experiments for the cubic nonlinear Schrödinger equation

In this chapter we conduct numerical experiments to confirm the results of Chapter 3 and 4 numerically. Observe that we have not analysed the space discretization error in Theorem 3.1, 4.1, 5.1 and 5.2, which comes into play in every numerical experiment. Therefore, we do not actually test these results in the following sections. By taking a small space mesh width, the results therein nevertheless give an indication whether our theorems are sharp or not.

We describe our general setting and our algorithm in Section 6.1. A crucial task is to generate discretized initial functions with a given regularity. We discuss our techniques how to gain these functions in Section 6.2. The correctness of our algorithm is confirmed by tests in Section 6.3. The most important question we address is whether the convergence order for initial functions with low regularity decreases in practice, see Theorem 4.1. We investigate this topic and additionally verify Theorem 3.1 in Section 6.4. Furthermore, the proofs of Theorem 3.1, 4.1, 5.1 and 5.2 suggest that the error increases when the H^4-, $H^{2+2\theta}$- or H^2-norm of the solution, respectively, increases. In the final Section 6.5 we use oscillating initial functions to confirm this conjecture.

6.1. An overview over the numerical experiments

The numerical computations are performed on the one-dimensional torus $\mathbb{T}^1$. We parametrize it by $[-\pi, \pi)$, discretize $[-\pi, \pi)$ by a uniform grid with 1024, 2048 or 4096 grid points, and equip them with periodic boundary conditions.

As explained in Section 2.2, the choice of the torus allows us to compute the solutions of the "linear" subproblem (2.6) in the Fourier space with the fast Fourier transform (FFT). The solutions of the subproblem (2.7) are obtained by a pointwise evaluation of the explicit solution formula (2.8).

Because we do not have explicit formulas for the solutions to problem (2.1), we have to calculate precise reference solutions. We conducted pre-experiments in which we computed them with the Strang splitting or with the forth order Yoshida scheme with very

small time step sizes. We obtained very similar results with both methods. So, we choose the Strang splitting for the computation of the reference solutions since it has the shorter computation times.

We choose $[0, 1]$ as time domain for our computations. This is possible since the solutions to (2.1) exist in one space dimension globally in time. The reference solutions are calculated with $R \cdot 2^7$ uniform time steps for all

$$\begin{aligned} R \in \{&128, 131, 134, 137, 140, 143, 146, 149, 152, 156, 159, 162, 166, 170, 173, 177, \\ &181, 185, 189, 193, 197, 202, 206, 211, 215, 220, 225, 230, 235, 240, 245, 251\}. \end{aligned}$$

The numbers that R takes as values are the rounded values of $\left(\sqrt[32]{2}\right)^k \cdot 128$ for $k = 0, \ldots, 31$. Thus, the time step sizes are almost uniformly distributed on a logarithmic scale. The solutions of the Strang splitting are computed with the numbers of uniform time steps that are the 128-th, 64-th, 32-th, 16-th and 8-th part of the ones for the reference solutions. All solutions, including the reference solutions, are saved at $R + 1$ equidistant time steps (including the starting time 0).

We measure the error of the Strang splitting by calculating at the $R + 1$ time points the discrete L^2-norm of the difference between the reference solution and the result of the Strang splitting computation. For this we choose that reference solution whose number of time steps is 2^l times the number of time steps of the Strang splitting for an $l \in \mathbb{N}$. The final errors are defined as the maximum over those discrete L^2-norms. We display them over the time step sizes in double logarithmic plots.

In order to illustrate the results of Chapter 3 and 4 numerically we would like to construct initial functions that are in certain Sobolev spaces H^s, but not in one with a higher order, i.e. not in H^r with $r > s$. Since we do not know how to do that, we construct initial functions that are in $H^{s-\varepsilon} \setminus H^s$ for all $\varepsilon \in (0, s)$. For shortness we say that such functions are *"almost in H^s"*. The arbitrary small difference between being in the certain H^s-space or not has no impact on numerical results. As regularity for the initial functions that are almost in H^s we choose $s = 4$, $7/2$, 3 and $5/2$.

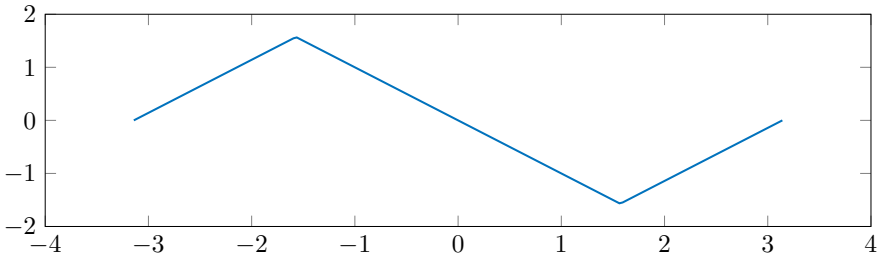
We use two different techniques to construct functions that are almost in H^s. The first one is to choose finitely many subintervals of $[-\pi, \pi)$ and a function that is smooth on each of these subintervals and given by an explicit formula. We then discretize these formulas on the space grid. For the second technique we use the Fourier representation of functions on the torus, draw randomly distributed Fourier coefficients and scale them appropriately. We describe both techniques in more detail in Section 6.2.

6.2. Construction of initial functions with a given regularity

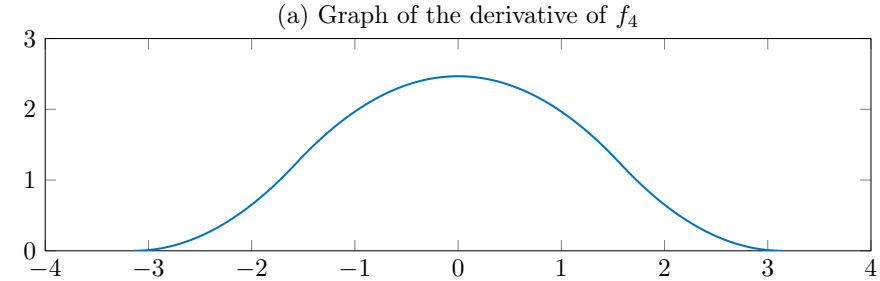
In this section we describe how we gain initial functions that are not in a certain H^s-space but in all larger Sobolev spaces, i.e. "almost in H^s". (Note that a larger Sobolev space has a smaller regularity parameter.) We use the technique of discretizing a function given by an explicit formula on the spatial grid and the one of drawing random Fourier coefficients. We normalize all gained initial functions in the discrete L^2-norm.

6.2.1. Discretising an explicitly given function

As basic functions for the construction of explicitly given initial functions we use piecewise linear functions and functions that are pieceweise of square root type. Piecewise linear functions are almost in $H^{3/2}$ and the square root is almost in H^1 (since their derivatives have Fourier coefficients of order $\frac{1}{k}$ and $\frac{1}{\sqrt{k}}$, respectively). By translating, mirroring and afterwards integrating (maybe more than once) we combine these basic functions to functions with the desired regularities. Observe that a function is continuous on the torus if and only if its canonical mapping onto the parametrization domain $[-\pi, \pi)$ satisfies periodic boundary conditions. So, the Sobolev embeddings give us constraints which of the following functions and derivatives have to satisfy periodic boundary conditions.



(a) Graph of the derivative of f_4



(b) Graph of f_4

Figure 6.1.: The graphs of the function f_4 and of its derivative.

We demonstrate the construction of the initial functions at the example of a function being almost in $H^{5/2}$. We choose the piecewise linear function whose graph is displayed

in Figure 6.1 (a). Integrating yields $f_4 : [-\pi, \pi) \to \mathbb{R}$ defined by

$$f_4(x) := \begin{cases} \frac{1}{2}(x+\pi)^2, & x \in [-\pi, -\frac{\pi}{2}), \\ -\frac{1}{2}x^2 + \frac{\pi^2}{4}, & x \in [-\frac{\pi}{2}, \frac{\pi}{2}), \\ \frac{1}{2}(x-\pi)^2, & x \in [\frac{\pi}{2}, \pi), \end{cases}$$

which is illustrated in Figure 6.1 (b).

In the same way we construct the function $f_1 : [-\pi, \pi) \to \mathbb{R}$ defined by

$$\begin{aligned}
&+\tfrac{8}{105}(x+\pi)^{7/2}, && x \in [-\pi, -\tfrac{7}{8}\pi), \\
&-\tfrac{8}{105}\big(-\big(x+\tfrac{6}{8}\pi\big)\big)^{7/2} + \tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}x^2 + \tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\tfrac{7}{8}\pi x \\
&\quad +\tfrac{16}{105}\big(\tfrac{\pi}{8}\big)^{7/2} + \tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\tfrac{7}{8}\pi\big)^2, && x \in [-\tfrac{7}{8}\pi, -\tfrac{6}{8}\pi), \\
&-\tfrac{8}{105}\big(\big(x+\tfrac{6}{8}\pi\big)\big)^{7/2} + \tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}x^2 + \tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\tfrac{7}{8}\pi x \\
&\quad +\tfrac{16}{105}\big(\tfrac{\pi}{8}\big)^{7/2} + \tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\tfrac{7}{8}\pi\big)^2, && x \in [-\tfrac{6}{8}\pi, -\tfrac{5}{8}\pi), \\
&+\tfrac{8}{105}\big(-\big(x+\tfrac{4}{8}\pi\big)\big)^{7/2} + \tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\tfrac{2}{8}\pi x \\
&\quad +\tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\big(\tfrac{5}{8}\pi\big)^2 - \big(\tfrac{7}{8}\pi\big)^2\big) + \tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\tfrac{10}{64}\pi^2 + \tfrac{14}{64}\pi^2\big), && x \in [-\tfrac{5}{8}\pi, -\tfrac{4}{8}\pi), \\
&-\tfrac{8}{105}\big(x+\tfrac{4}{8}\pi\big)^{7/2} + \tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\tfrac{2}{8}\pi x \\
&\quad +\tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\big(\tfrac{5}{8}\pi\big)^2 - \big(\tfrac{7}{8}\pi\big)^2\big) + \tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\tfrac{10}{64}\pi^2 + \tfrac{14}{64}\pi^2\big), && x \in [-\tfrac{4}{8}\pi, -\tfrac{3}{8}\pi), \\
&+\tfrac{8}{105}\big(-\big(x+\tfrac{2}{8}\pi\big)\big)^{7/2} - \tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}x^2 - \tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\tfrac{\pi}{8}x - \tfrac{16}{105}\big(\tfrac{\pi}{8}\big)^{7/2} \\
&\quad +\tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\big(\tfrac{3}{8}\pi\big)^2 + \big(\tfrac{5}{8}\pi\big)^2 - \big(\tfrac{7}{8}\pi\big)^2\big) \\
&\quad +\tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(-\tfrac{3}{64}\pi^2 - \tfrac{6}{64}\pi^2 + \tfrac{10}{64}\pi^2 + \tfrac{14}{64}\pi^2\big), && x \in [-\tfrac{3}{8}\pi, -\tfrac{2}{8}\pi), \\
&+\tfrac{8}{105}(x+\tfrac{2}{8}\pi)^{7/2} - \tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}x^2 - \tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\tfrac{\pi}{8}x - \tfrac{16}{105}\big(\tfrac{\pi}{8}\big)^{7/2} \\
&\quad +\tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\big(\tfrac{3}{8}\pi\big)^2 + \big(\tfrac{5}{8}\pi\big)^2 - \big(\tfrac{7}{8}\pi\big)^2\big) \\
&\quad +\tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(-\tfrac{3}{64}\pi^2 - \tfrac{6}{64}\pi^2 + \tfrac{10}{64}\pi^2 + \tfrac{14}{64}\pi^2\big), && x \in [-\tfrac{2}{8}\pi, -\tfrac{\pi}{8}), \\
&-\tfrac{8}{105}(-x)^{7/2} + \tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(-\big(\tfrac{\pi}{8}\big)^2 + \big(\tfrac{3}{8}\pi\big)^2 + \big(\tfrac{5}{8}\pi\big)^2 - \big(\tfrac{7}{8}\pi\big)^2\big) \\
&\quad +\tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\tfrac{1}{64}\pi^2 - \tfrac{3}{64}\pi^2 - \tfrac{6}{64}\pi^2 + \tfrac{10}{64}\pi^2 + \tfrac{14}{64}\pi^2\big), && x \in [-\tfrac{\pi}{8}, 0), \\
&-\tfrac{8}{105}x^{7/2} + \tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(-\big(\tfrac{\pi}{8}\big)^2 + \big(\tfrac{3}{8}\pi\big)^2 + \big(\tfrac{5}{8}\pi\big)^2 - \big(\tfrac{7}{8}\pi\big)^2\big) \\
&\quad +\tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\tfrac{1}{64}\pi^2 - \tfrac{3}{64}\pi^2 - \tfrac{6}{64}\pi^2 + \tfrac{10}{64}\pi^2 + \tfrac{14}{64}\pi^2\big), && x \in [0, \tfrac{\pi}{8}), \\
&+\tfrac{8}{105}\big(-\big(x-\tfrac{2}{8}\pi\big)\big)^{7/2} - \tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}x^2 + \tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\tfrac{\pi}{8}x - \tfrac{16}{105}\big(\tfrac{\pi}{8}\big)^{7/2} \\
&\quad +\tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\big(\tfrac{3}{8}\pi\big)^2 + \big(\tfrac{5}{8}\pi\big)^2 - \big(\tfrac{7}{8}\pi\big)^2\big) \\
&\quad +\tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(-\tfrac{3}{64}\pi^2 - \tfrac{6}{64}\pi^2 + \tfrac{10}{64}\pi^2 + \tfrac{14}{64}\pi^2\big), && x \in [\tfrac{\pi}{8}, \tfrac{2}{8}\pi), \\
&+\tfrac{8}{105}\big(\big(x-\tfrac{2}{8}\pi\big)\big)^{7/2} - \tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}x^2 + \tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\tfrac{\pi}{8}x - \tfrac{16}{105}\big(\tfrac{\pi}{8}\big)^{7/2}, \\
&\quad +\tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\big(\tfrac{3}{8}\pi\big)^2 + \big(\tfrac{5}{8}\pi\big)^2 - \big(\tfrac{7}{8}\pi\big)^2\big) \\
&\quad +\tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(-\tfrac{3}{64}\pi^2 - \tfrac{6}{64}\pi^2 + \tfrac{10}{64}\pi^2 + \tfrac{14}{64}\pi^2\big), && x \in [\tfrac{2}{8}\pi, \tfrac{3}{8}\pi), \\
&-\tfrac{8}{105}\big(-\big(x-\tfrac{4}{8}\pi\big)\big)^{7/2} - \tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\tfrac{2}{8}\pi x \\
&\quad +\tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\big(\tfrac{5}{8}\pi\big)^2 - \big(\tfrac{7}{8}\pi\big)^2\big) + \tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\tfrac{10}{64}\pi^2 + \tfrac{14}{64}\pi^2\big), && x \in [\tfrac{3}{8}\pi, \tfrac{4}{8}\pi), \\
&+\tfrac{8}{105}\big(x-\tfrac{4}{8}\pi\big)^{7/2} - \tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\tfrac{2}{8}\pi x \\
&\quad +\tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\big(\tfrac{5}{8}\pi\big)^2 - \big(\tfrac{7}{8}\pi\big)^2\big) + \tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\tfrac{10}{64}\pi^2 + \tfrac{14}{64}\pi^2\big), && x \in [\tfrac{4}{8}\pi, \tfrac{5}{8}\pi),
\end{aligned}$$

$$
\begin{aligned}
&-\tfrac{8}{105}\big(-\big(x-\tfrac{6}{8}\pi\big)\big)^{7/2}+\tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}x^2-\tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\tfrac{7}{8}\pi x\\
&\quad+\tfrac{16}{105}\big(\tfrac{\pi}{8}\big)^{7/2}+\tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\tfrac{7}{8}\pi\big)^2, && x\in[\tfrac{5}{8}\pi,\tfrac{6}{8}\pi),\\
&-\tfrac{8}{105}\big(x-\tfrac{6}{8}\pi\big)^{7/2}+\tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}x^2-\tfrac{4}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\tfrac{7}{8}\pi x\\
&\quad+\tfrac{16}{105}\big(\tfrac{\pi}{8}\big)^{7/2}+\tfrac{2}{3}\big(\tfrac{\pi}{8}\big)^{3/2}\big(\tfrac{7}{8}\pi\big)^2, && x\in[\tfrac{6}{8}\pi,\tfrac{7}{8}\pi),\\
&+\tfrac{8}{105}\big(-\big(x-\pi\big)\big)^{7/2}, && x\in[\tfrac{7}{8}\pi,\pi),
\end{aligned}
$$

the function $f_2:[-\pi,\pi)\to\mathbb{R}$ defined by

$$
f_2(x):=\begin{cases}
\frac{1}{6}(x+\pi)^3, & x\in[-\pi,-\frac{3}{4}\pi),\\
-\frac{1}{6}(x+\frac{\pi}{2})^3+\frac{\pi^2}{16}x+\frac{3\pi^3}{64}, & x\in[-\frac{3}{4}\pi,-\frac{\pi}{4}),\\
\frac{1}{6}x^3+\frac{\pi^3}{32}, & x\in[-\frac{\pi}{4},0),\\
-\frac{1}{6}x^3+\frac{\pi^3}{32}, & x\in[0,\frac{\pi}{4}),\\
\frac{1}{6}(x-\frac{\pi}{2})^3+\frac{\pi^2}{16}x-\frac{3\pi^3}{64}, & x\in[\frac{\pi}{4},\frac{3}{4}\pi),\\
-\frac{1}{6}(x-\pi)^3, & x\in[\frac{3}{4}\pi,\pi),
\end{cases}
$$

and the function $f_3:[-\pi,\pi)\to\mathbb{R}$ defined by

$$
f_3(x):=\begin{cases}
\frac{4}{15}(x+\pi)^{5/2}, & x\in[-\pi,-\frac{3}{4}\pi),\\
\frac{4}{15}\big(-(x+\frac{\pi}{2})\big)^{5/2}+\frac{4}{3}\big(\frac{\pi}{4}\big)^{3/2}x+\frac{4}{3}\big(\frac{\pi}{4}\big)^{3/2}\frac{3}{4}\pi, & x\in[-\frac{3}{4}\pi,-\frac{1}{2}\pi),\\
-\frac{4}{15}\big(x+\frac{\pi}{2}\big)^{5/2}+\frac{4}{3}\big(\frac{\pi}{4}\big)^{3/2}x+\frac{4}{3}\big(\frac{\pi}{4}\big)^{3/2}\frac{3}{4}\pi, & x\in[-\frac{\pi}{2},-\frac{\pi}{4}),\\
-\frac{4}{15}(-x)^{5/2}+\frac{4}{3}\big(\frac{\pi}{4}\big)^{3/2}\frac{\pi}{2}, & x\in[-\frac{\pi}{4},0),\\
-\frac{4}{15}x^{5/2}+\frac{4}{3}\big(\frac{\pi}{4}\big)^{3/2}\frac{\pi}{2}, & x\in[0,\frac{\pi}{4}),\\
-\frac{4}{15}\big(-(x-\frac{\pi}{2})\big)^{5/2}-\frac{4}{3}\big(\frac{\pi}{4}\big)^{3/2}x+\frac{4}{3}\big(\frac{\pi}{4}\big)^{3/2}\frac{3}{4}\pi, & x\in[\frac{\pi}{4},\frac{\pi}{2}),\\
\frac{4}{15}\big(x-\frac{\pi}{2}\big)^{5/2}-\frac{4}{3}\big(\frac{\pi}{4}\big)^{3/2}x+\frac{4}{3}\big(\frac{\pi}{4}\big)^{3/2}\frac{3}{4}\pi, & x\in[\frac{1}{2}\pi,\frac{3}{4}\pi),\\
\frac{4}{15}(-(x-\pi))^{5/2}, & x\in[\frac{3}{4}\pi,\pi).
\end{cases}
$$

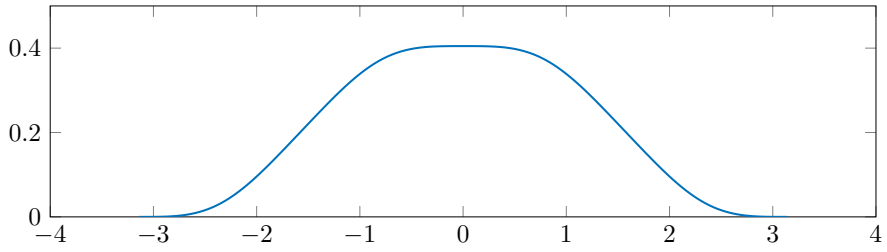
The graphs of these functions are displayed in Figure 6.2.

We claim that f_1 is almost in H^4, f_2 almost in $H^{7/2}$, f_3 almost in H^3 and f_4 almost in $H^{5/2}$. The proof for f_1 can be done in the same way as the following ones for the other functions.

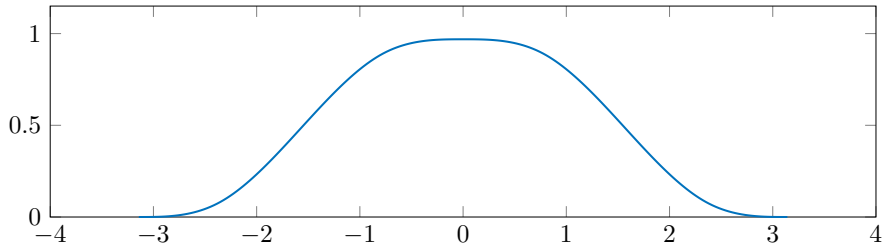
For the Fourier coefficients $\{c_k, k\in\mathbb{Z}\}$ for f_2 we get for all $k\in\mathbb{Z}\setminus\{0\}$ from a long calculation, using integration by parts, that

$$
c_k=\frac{1}{k^4\sqrt{2\pi}}\big(e^{-ik\pi}-2e^{-ik\frac{3}{4}\pi}+2e^{-ik\frac{1}{4}\pi}+2e^{ik\frac{1}{4}\pi}-2e^{ik\frac{3}{4}\pi}+e^{ik\pi}\big).
$$

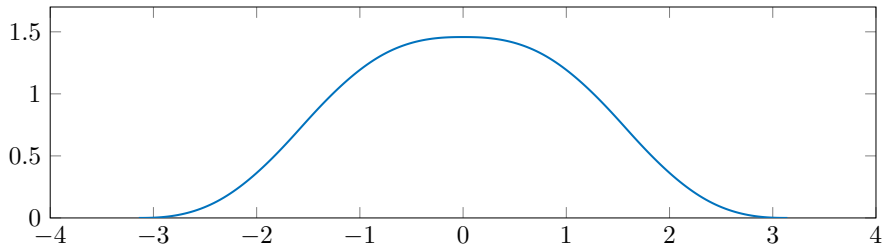
This shows the desired regularity since the series $\sum_{k\in\mathbb{Z}}\frac{1}{k^\alpha}$ is convergent if and only if $\alpha>1$.



(a) Graph of f_1



(b) Graph of f_2



(c) Graph of f_3

Figure 6.2.: The graphs of the functions f_1, f_2 and f_3.

For the Fourier coefficients $\{c_k, k \in \mathbb{Z}\}$ of f_3 we first compute in a similar way for all $k \in \mathbb{Z} \setminus \{0\}$ the identity

$$c_k = \frac{1}{-k^2\sqrt{2\pi}}\Bigg(\left(e^{-ik\pi} - e^{-ik\pi/2} - 1 + e^{ik\pi/2}\right)\int_0^{\pi/4} x^{1/2}e^{ikx}\,\mathrm{d}x + \left(e^{-ik\pi/2} - 1 - e^{ik\pi/2} + e^{ik\pi}\right)\int_0^{\pi/4} x^{1/2}e^{-ikx}\,\mathrm{d}x\Bigg).$$

We have with the substitution $x = y^2$ that

$$\begin{aligned}\int_0^{\pi/4} x^{1/2}e^{ikx}\,\mathrm{d}x &= \int_0^{\sqrt{\pi/4}} 2y^2e^{iky^2}\,\mathrm{d}y \\ &= -\int_0^{\sqrt{\pi/4}} \frac{1}{ik}e^{iky^2}\,\mathrm{d}y + \left[\frac{1}{ik}ye^{iky^2}\right]_{y=0}^{\sqrt{\pi/4}}.\end{aligned}$$

Assuming without loss of generality that $k > 0$, we get with the substitution $z = y\sqrt{k}$ that

$$\int_0^{\pi/4} x^{1/2}e^{ikx}\,\mathrm{d}x = -\int_0^{\sqrt{k\pi/4}} \frac{1}{ik^{3/2}}e^{iz^2}\,\mathrm{d}z + \frac{1}{ik}\sqrt{\frac{\pi}{4}}e^{ik\pi/4}.$$

From

$$\overline{\int_0^{\pi/4} x^{1/2} e^{-ikx}\,\mathrm{d}x} = \int_0^{\pi/4} x^{1/2} e^{ikx}\,\mathrm{d}x$$

we thus deduce

$$c_k = \frac{1}{-ik^{7/2}\sqrt{2\pi}}\Bigg(\left(-e^{-ik\pi} + e^{-ik\pi/2} + 1 - e^{ik\pi/2}\right)\int_0^{\sqrt{k\pi/4}} e^{iz^2}\,\mathrm{d}z + \left(e^{-ik\pi/2} - 1 - e^{ik\pi/2} + e^{ik\pi}\right)\int_0^{\sqrt{k\pi/4}} e^{-iz^2}\,\mathrm{d}z\Bigg).$$

For all $k \geq 0$ we get with the substitutions $y = z^2$ and $y = z^2 - \pi$ that

$$\begin{aligned}\int_{\sqrt{2k\pi}}^{\sqrt{2(k+1)\pi}} \sin(z^2)\,\mathrm{d}z &= \int_{\sqrt{2k\pi}}^{\sqrt{(2k+1)\pi}} \sin(z^2)\,\mathrm{d}z + \int_{\sqrt{(2k+1)\pi}}^{\sqrt{2(k+1)\pi}} \sin(z^2)\,\mathrm{d}z \\ &= \frac{1}{2}\int_{2k\pi}^{(2k+1)\pi} \left(\tfrac{1}{\sqrt{y}} - \tfrac{1}{\sqrt{y+\pi}}\right)\sin(y)\,\mathrm{d}y > 0. \end{aligned} \tag{6.1}$$

Moreover, the function $t \mapsto \int_{\sqrt{2k\pi}}^{t} \sin(z^2)\,\mathrm{d}z$ is increasing on $\left[\sqrt{2k\pi}, \sqrt{(2k+1)\pi}\right]$ and decreasing on $\left[\sqrt{(2k+1)\pi}, \sqrt{2(k+1)\pi}\right]$ for all $k \geq 0$. Together this gives

$$\begin{aligned}\inf_{k\in\mathbb{N}} \left|\int_0^{\sqrt{k\pi/4}} e^{iz^2}\,\mathrm{d}z\right| &\geq \inf_{t\geq\sqrt{\pi/4}} \left|\int_0^t e^{iz^2}\,\mathrm{d}z\right| \geq \inf_{t\geq\sqrt{\pi/4}} \left|\int_0^t \sin(z^2)\,\mathrm{d}z\right| \\ &= \min\left\{\left|\int_0^{\sqrt{\pi/4}} \sin(z^2)\,\mathrm{d}z\right|, \left|\int_0^{\sqrt{2\pi}} \sin(z^2)\,\mathrm{d}z\right|\right\} > 0.\end{aligned}$$

An analogous calculation as (6.1) gives

$$\int_{\sqrt{(2k+1)\pi}}^{\sqrt{(2k+3)\pi}} \sin(z^2)\,\mathrm{d}z < 0$$

for all $k \geq 0$. Together with (6.1) this yields that the non-negative sequence $(b_l)_{l\in\mathbb{N}}$ being defined by $b_l := \left|\int_{\sqrt{l\pi}}^{\sqrt{(l+1)\pi}} \sin(z^2)\,\mathrm{d}z\right|$ for all $l \in \mathbb{N}$ is monotonically decreasing. The computation

$$b_l \leq \sqrt{(l+1)\pi} - \sqrt{l\pi} \leq \sup_{t\in[l\pi,(l+1)\pi]} \frac{1}{2\sqrt{t}} \cdot \pi = \frac{\sqrt{\pi}}{2\sqrt{l}} \longrightarrow 0$$

as $l \to \infty$ shows that (b_l) is a null sequence. Therefore, the Leibniz test ensures

$$\int_0^{\infty} \sin(z^2)\,\mathrm{d}z = \sum_{l=0}^{\infty} \int_{\sqrt{l\pi}}^{\sqrt{(l+1)\pi}} \sin(z^2)\,\mathrm{d}z < \infty.$$

With an analogous calculation for the cosine we see that

$$\begin{aligned}\sup_{k\in\mathbb{N}}\left|\int_0^{\sqrt{k\pi/4}} e^{iz^2}\,\mathrm{d}z\right| &\leq \sup_{t\geq 0}\left|\int_0^t e^{iz^2}\,\mathrm{d}z\right| \\ &\leq \sup_{t\geq 0}\left|\int_0^t \cos(z^2)\,\mathrm{d}z\right| + \sup_{t\geq 0}\left|\int_0^t \sin(z^2)\,\mathrm{d}z\right| \\ &\leq \int_0^{\sqrt{\pi/2}} \cos(z^2)\,\mathrm{d}z - \int_{\sqrt{\pi/2}}^{\infty} \cos(z^2)\,\mathrm{d}z + \int_0^{\infty} \sin(z^2)\,\mathrm{d}z < \infty.\end{aligned}$$

Hence, we have constants $C_1, C_2 > 0$ such that

$$C_1 \leq \left|\int_0^{\sqrt{k\pi/4}} e^{iz^2}\,\mathrm{d}z\right| \leq C_2$$

for all $k \in \mathbb{N}$. This finishes the regularity proof for the same reason as above.

For the Fourier coefficients $\{c_k, k \in \mathbb{Z}\}$ of f_4 we get for all $k \in \mathbb{Z} \setminus \{0\}$ with a similar but shorter calculation than the one for f_2 that

$$c_k = \frac{2}{-ik^3\sqrt{2\pi}}\left(e^{-ik\pi/2} - e^{ik\pi/2}\right).$$

Analogously to the argumentation for f_2 the claim follows.

6.2.2. Drawing randomly distributed Fourier coefficients

Another technique for gaining initial functions being almost in H^s is to use the representation of the Sobolev spaces via the Fourier transform, see (1.1). We work with two variants of this idea. The first one is to use N Fourier coefficients $c_{-N/2}, \ldots, c_{N/2-1}$ that are drawn with a normally distributed real part and a normally distributed imaginary part. The coefficients are scaled by multiplying them with $\left(1 + |\xi|^2\right)^{s/2}$, where ξ is the variable in Fourier space and s the degree of "regularity" of the function. Afterwards, the inverse FFT is applied to get the values of the function on the space grid. The second idea is to draw an angle φ_k from a uniform distribution on $[0, 2\pi)$, to set the Fourier coefficient c_k to $\exp(i\varphi_k)$ for all $k \in \{-N/2, \ldots, N/2 - 1\}$ and also to apply the inverse FFT.

6.3. Testing of the Strang splitting scheme

In this section we test our numerical programme to confirm its correctness. We do this by computing the numerical approximation to plane wave solutions and to mollified soliton solutions. In this section we discretize $[-\pi, \pi)$ with $N = 1024$ space grid points.

6.3.1. Plane wave solutions

The plane wave

$$u(t,x) = a\exp(ix)\exp(-it)\exp(a^2 i\mu t)$$

for $x \in \mathbb{R}$ with parameter $a \in \mathbb{R}$ and initial function

$$u_0(x) = u(0,x) = a\exp(ix)$$

for $x \in \mathbb{R}$ is a 2π-periodic solution to (2.5) on the full space $\mathbb{R}$. Restricting it to $[-\pi,\pi)$ and mapping this restriction to the torus $\mathbb{T}$ via its parametrization gives a solution to (2.5) on $\mathbb{T}$. For the following experiments we choose the parameter a such that u_0 has norm 1 in the discrete L^2-norm. Since we have an explicit formula for the solution, we do not need to compute a reference solution.

Figure 6.3 shows errors of a very small magnitude for both the defocusing and the focusing case. This could be expected since the action of the solution (2.8) to the "nonlinear" subproblem (2.7) on the exact solution u at an arbitrary time point is only the multiplication with a constant depending on the time step size.

The errors that we see are maybe the result of rounding errors. This conjecture is supported by the fact that the errors are higher for smaller time step sizes, rising approximately with order one in the number of time steps.

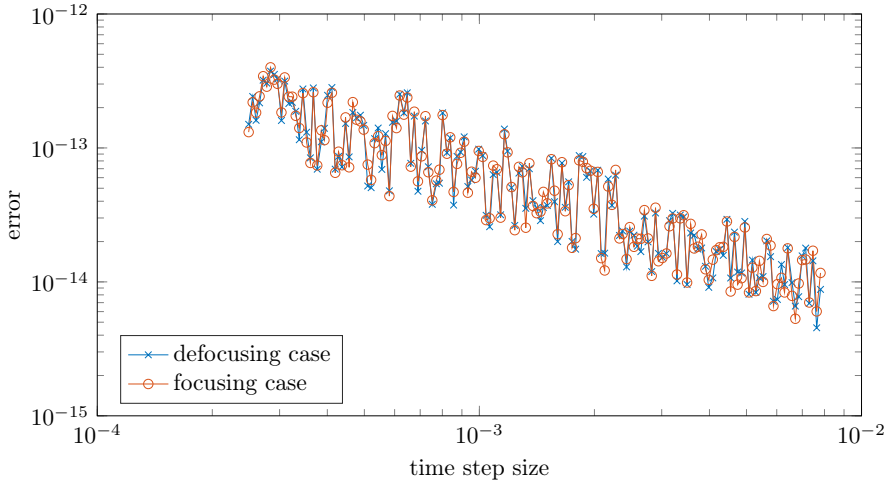


Figure 6.3.: Error of the Strang splitting for a plane wave solution.

6.3.2. Soliton solutions

The soliton

$$u(t,x) = \frac{a\sqrt{2}}{\cosh(ax)}\exp(a^2 it) \tag{6.2}$$

for $x \in \mathbb{R}$ with parameter $a \in \mathbb{R}$ and initial function

$$u_0(x) = u(0,x) = \frac{a\sqrt{2}}{\cosh(ax)}$$

for $x \in \mathbb{R}$ is a solution to (2.5) in the focusing case ($\mu = -1$) on the full space $\mathbb{R}$. If we restrict u_0 to the parametrization interval $[-\pi, \pi)$ and identify that with $\mathbb{T}$, we see that the restricted initial function is not differentiable on the torus since it has a kink at that point of the torus that is identified with the point $-\pi$ of the parametrization interval. Therefore, we first discretize the standard mollifier

$$\psi(x) = \begin{cases} \exp\left(-\frac{1}{1-(10x/\pi)^2}\right), & x \in \left[-\frac{\pi}{10}, \frac{\pi}{10}\right], \\ 0, & x \in [-\pi, \pi) \setminus \left[-\frac{\pi}{10}, \frac{\pi}{10}\right], \end{cases}$$

on the space grid and normalize it in the discrete L^2-norm. Then we convolute it with the restricted u_0. As always, we normalize the resulting function in the discrete L^2-norm.

We choose $a_1 = 5/2$ and $a_2 = 4$ for the following experiments. The parameter a_1 leads to a soliton with a broad peak and the parameter a_2 to one with a narrow peak. Figure 6.4 shows the results of the computation. We clearly see the convergence order two. The error is larger for the narrow soliton. This is maybe caused by the fact that the space grid resolves the thin peak worse than the peak of the broad soliton.

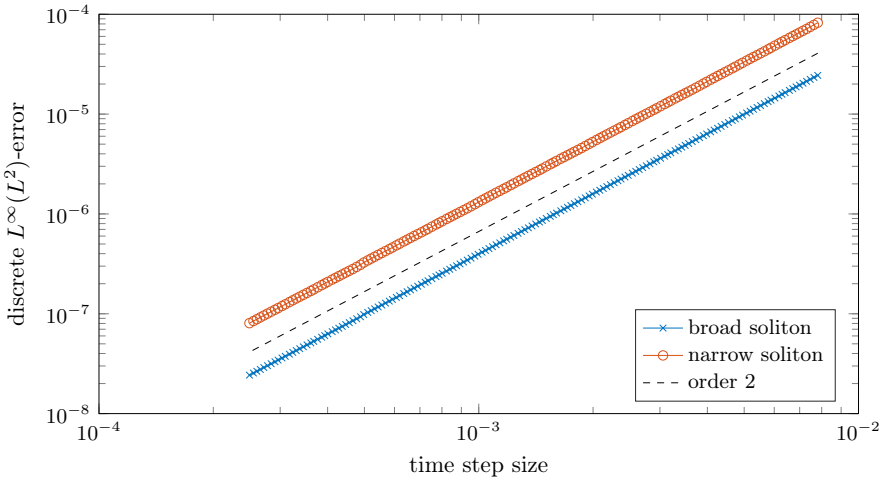


Figure 6.4.: Error of the Strang splitting for two mollified soliton solutions.

The smoothing of the initial function is necessary for the well-definedness of the algorithm, which we see by the following experiment. We do neither convolute the initial function with a mollifier nor normalize it in the discrete L^2-norm. By a very fine resolution of a small part of the time step size range we get Figure 6.5, which shows very large errors for some particular time step sizes. For the sake of comparison we do not use the exact solution formula, see (6.2), as reference solution, and we additionally display the errors we gain for the (also not normalized) mollified initial function.

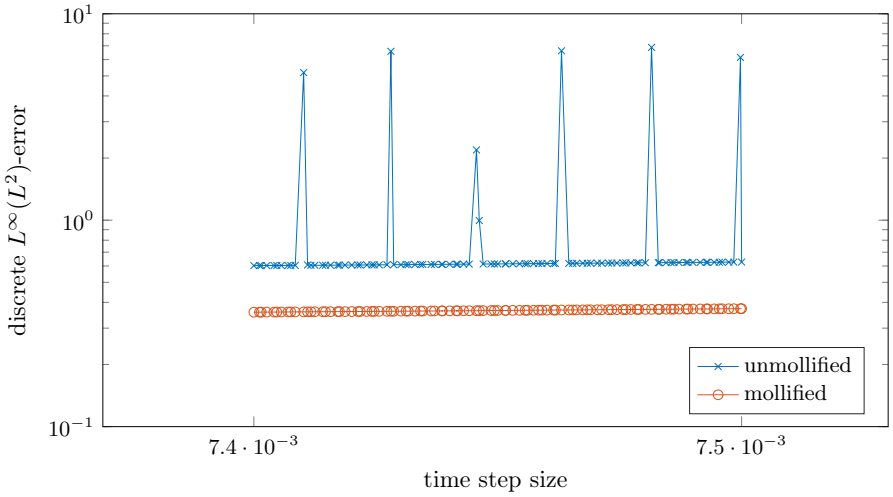


Figure 6.5.: Comparison of the errors of the Strang splitting for a mollified and an unmollified version of a soliton solution.

6.4. Convergence orders of the Strang splitting scheme

In the experiments in this section we investigate the convergence order of the Strang splitting scheme. We first confirm the second order convergence for initial functions in H^4, see Theorem 3.1. Afterwards we want to find out whether the convergence order of the Strang splitting is reduced in the case that the initial functions are not in H^4 but only in an H^s with $s \in (2, 4)$, see Theorem 4.1.

6.4.1. Results of the experiments with initial functions in H^4

We confirm the second order convergence for initial functions in H^4 with initial functions that are almost in H^4. The results of the computations are displayed in Figures 6.6 and 6.7. In both diagrams we see clearly a convergence order of two. There are only very few time step sizes where the error is larger than expected from the other values.

6.4.2. Results of the experiments with less regular initial functions

We continue with experiments with initial functions being almost in $H^{7/2}$, H^3 and $H^{5/2}$. The results can be seen in Figures 6.8, 6.9, 6.10, 6.11, 6.12 and 6.13. The diagrams show an oscillating behaviour of the error. We can only speculate about the reasons for this. Two possible explanations are the following ones.

The first one is that the data points with the higher magnitudes were disturbed by resonance effects in the computations. This would lead to the conclusion that the convergence

order is two also in this case and that no order reduction can be seen. This might be due to the fact that there are functions that would show the sharpness of Theorem 4.1, but that the ones we have chosen for our experiments are better in the sense that their error behaves like the one we would expect of a function of higher regularity. Of course it is impossible to make experiments with all Sobolev functions of (almost) a given regularity, but maybe a more clever choice of the initial functions can reveal an order reduction. The other possibility is that the convergence orders obtained in Theorem 4.1 (and Theorem 5.1) are too pessimistic. Maybe one can see with another proof strategy convergence of a higher order, perhaps even one with the classical order two.

The second possible explanation is that for many time step sizes the error is, due to cancellation effects in the computations, smaller than expected. Then the diagrams show a reduction of the convergence order of almost the amount that Theorem 4.1 predicts.

It is remarkable that the oscillations occur at more time step sizes and are much higher if we use the initial functions gained by drawing Fourier coefficients than if we use the ones stemming from an explicit formula. The reason is maybe the amount of points that hinder the initial function from being in a higher-order Sobolev space. For the function with explicit formula it consists of the finitely many boundary points of the parts of its domain and is thus a Lebesgue null set. In contrast to this the functions from the randomly drawn (uniformly or normally distributed) Fourier coefficient are, in the limit of the number of space grid points going to infinity, of the low regularity on every open subset of the domain $[-\pi, \pi)$.

The different magnitudes of the errors for one and the same time step size are caused by different values of the error constant. Due to Theorems 3.1 and 4.1 the error constant depends on the supremum of the H^s-norms of the exact solutions. We approximate these suprema by the fully discrete $L^\infty(H^s)$-norm of that corresponding reference solution with the smallest time step size. For the case of explicitly given initial functions we get the values

	H^4	$H^{7/2}$	H^3	$H^{5/2}$
$N = 1024$	10.0616	5.4524	2.6762	1.9731
$N = 2048$	10.3906	5.6551	2.7251	2.0033
$N = 4096$	10.7186	5.8506	2.7732	2.0332

in the defocusing and the values

	H^4	$H^{7/2}$	H^3	$H^{5/2}$
$N = 1024$	10.2816	5.5240	2.6569	1.9124
$N = 2048$	10.6036	5.7241	2.7062	1.9436
$N = 4096$	10.9180	5.9174	2.7546	1.9743

in the focusing case. For the case of initial functions gained by normally distributed

Fourier coefficients we obtain the values

	H^4	$H^{7/2}$	H^3	$H^{5/2}$
$N = 1024$	37.4775	34.9033	31.9712	28.7382
$N = 2048$	81.0564	70.8092	61.0943	52.0103
$N = 4096$	65.8371	64.4928	62.6448	60.0718

in the defocusing and the values

	H^4	$H^{7/2}$	H^3	$H^{5/2}$
$N = 1024$	37.5064	34.9237	31.9849	28.7472
$N = 2048$	81.0797	70.8239	61.1033	52.0158
$N = 4096$	65.8278	64.4926	62.6441	60.0710

in the focusing case. For the case of initial functions gained by uniformly distributed Fourier coefficients we obtain the values

	H^4	$H^{7/2}$	H^3	$H^{5/2}$
$N = 1024$	30.1452	29.4183	28.4263	27.0721
$N = 2048$	42.6113	41.5857	40.1860	38.2741
$N = 4096$	60.2558	58.8014	56.8236	54.1214

in the defocusing and the values

	H^4	$H^{7/2}$	H^3	$H^{5/2}$
$N = 1024$	30.1260	29.4018	28.4130	27.0619
$N = 2048$	42.6081	41.5837	40.1848	38.2736
$N = 4096$	60.2565	58.8033	56.8255	54.1236

in the focusing case. Comparing these values explains why the errors belonging to the case of normally distributed Fourier coefficients are, relatively to the errors belonging to the other two cases, for the choice $N = 2048$ larger than for the other two space discretizations, and why this effect is weaker for less regular initial functions. Furthermore, it explains why the errors belonging to the explicitly given initial functions are, relatively to the errors belonging from the other types of initial functions, smaller for less regular initial functions.

6.5. Increase of the error constant for highly oscillating initial functions

The proofs of Theorems 3.1, 4.1, 5.1 and 5.2 show that the error constant increases if the supremum of the H^4-norm, the $H^{2+2\theta}$-norm or the H^2-norm of the solution, respectively,

enlarges. Because we cannot control the norms of the solutions itself, we adjust the norms of the initial functions.

As initial functions for the following experiment we use for the factors $K \in \{1, 2, 4, 8\}$ the smooth functions

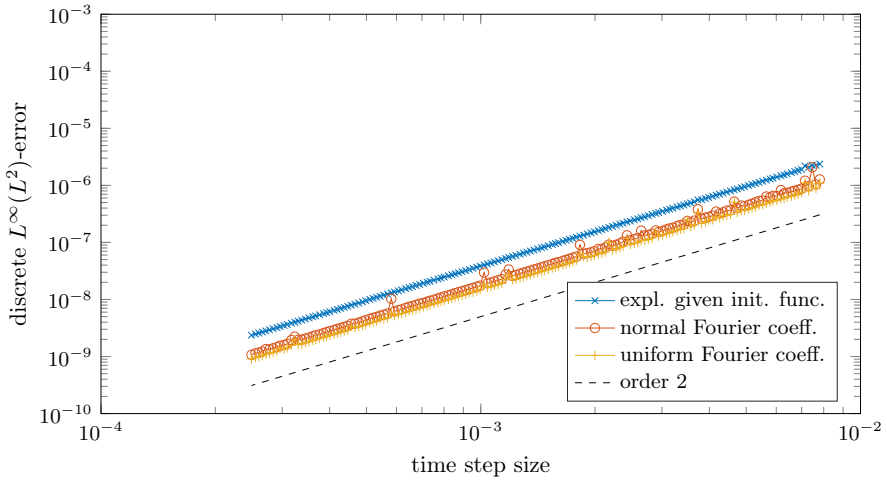
$$x \mapsto \sin(Kx) + \cos((K+1)x)$$

and normalize them in the discrete L^2-norm after the discretization on $N = 1024$ equidistant space grid points. They have an increasing H^4-norm, but are not just scalings of one another with a different oscillation frequency. The latter fact has the advantage that we have slightly different "types" of oscillating functions, so that the results of the following calculations are probably not caused by a similar structure of the initial functions and the solutions. We see clearly that the error is larger when the initial function is more rapidly oscillating, see Figure 6.14.

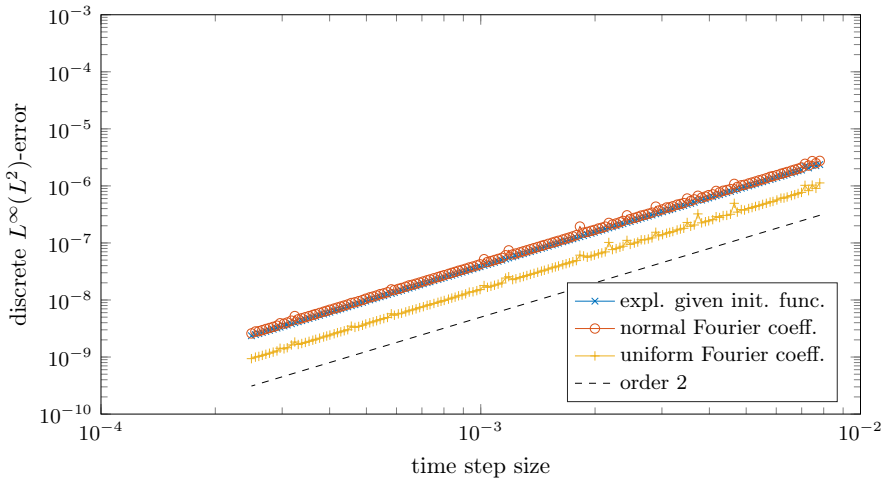
As in Section 6.4, we use the fully discrete $L^\infty(H^4)$-norm of the reference solutions with the smallest time step sizes as an approximation to the $L^\infty(H^4)$-norm of the exact solution. The resulting values

K	1	2	4	8
defocusing case	19.2	75.0	524.6	5629.2
focusing case	19.1	74.9	523.0	5621.0

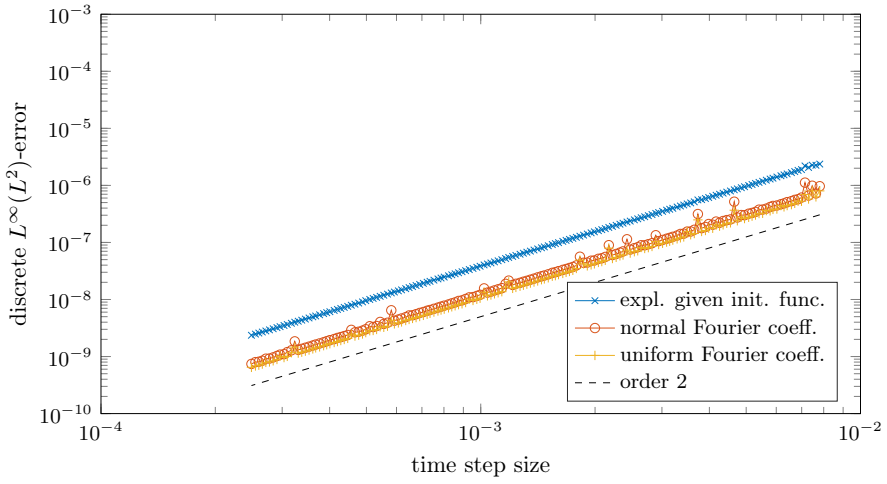
explain the increase of the errors for K increasing.



(a) $N = 1024$.

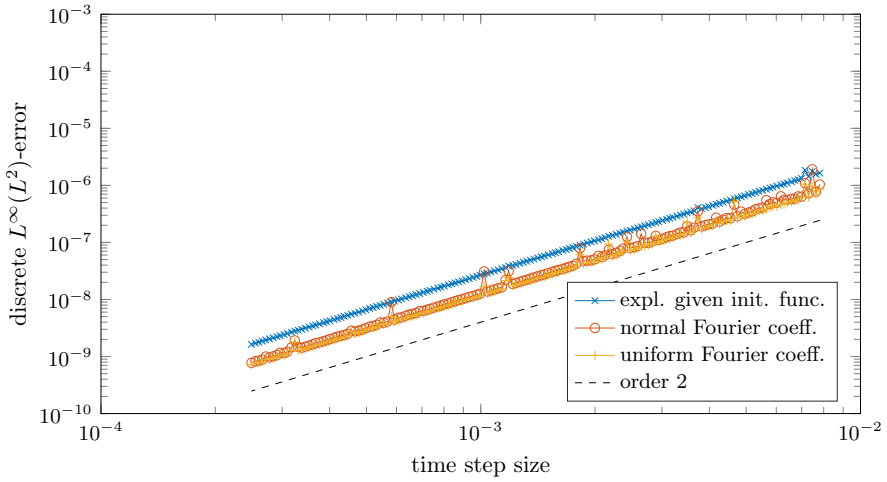


(b) $N = 2048$.

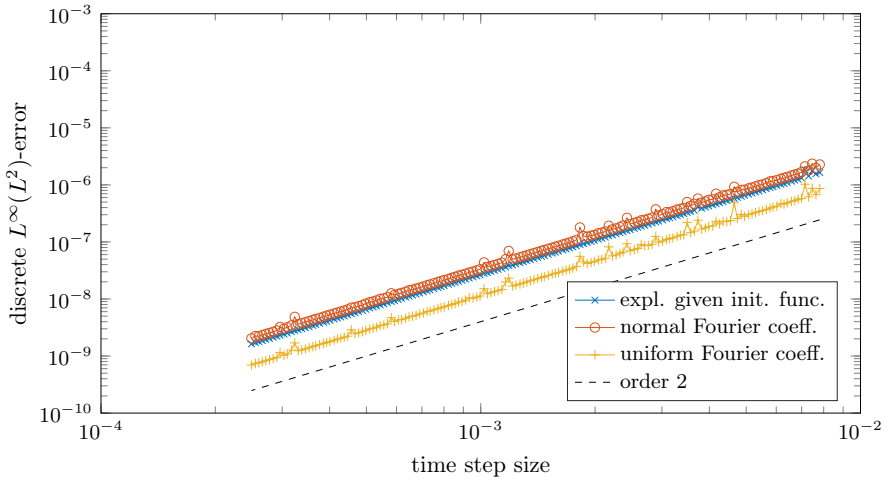


(c) $N = 4096$.

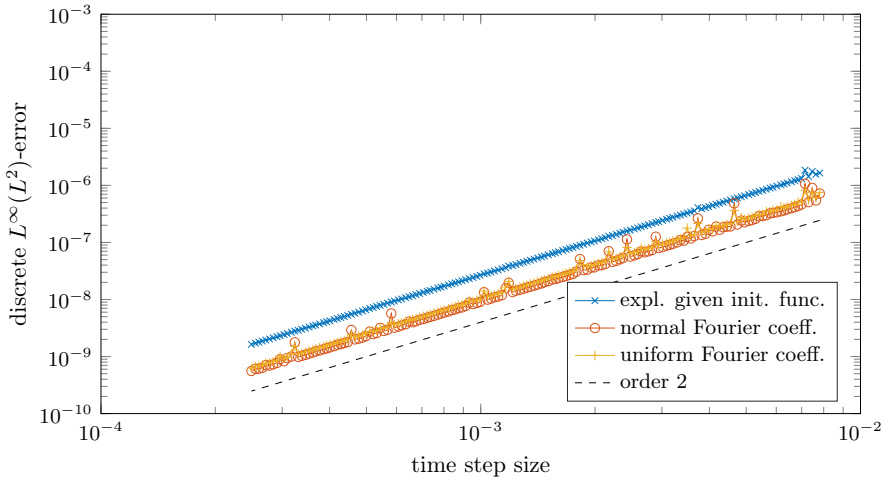
Figure 6.6.: Errors of the Strang splitting for initial functions being almost in H^4 in the defocusing case for N space grid points.



(a) $N = 1024$.

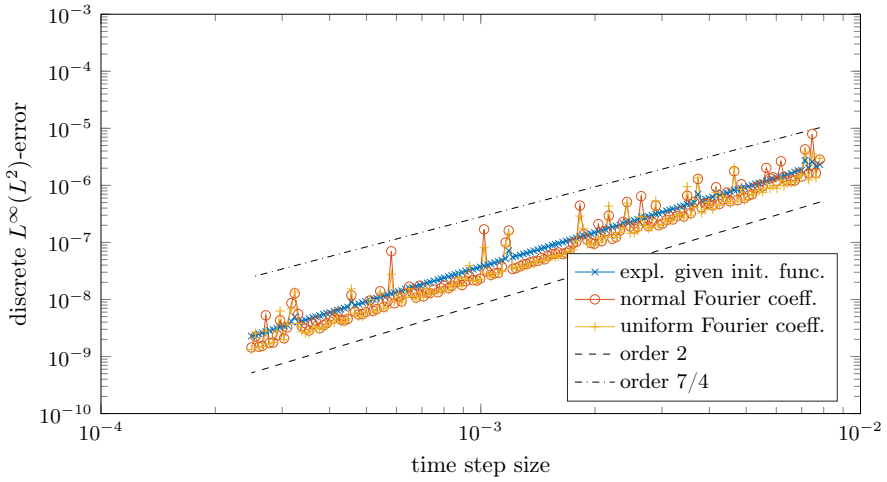


(b) $N = 2048$.

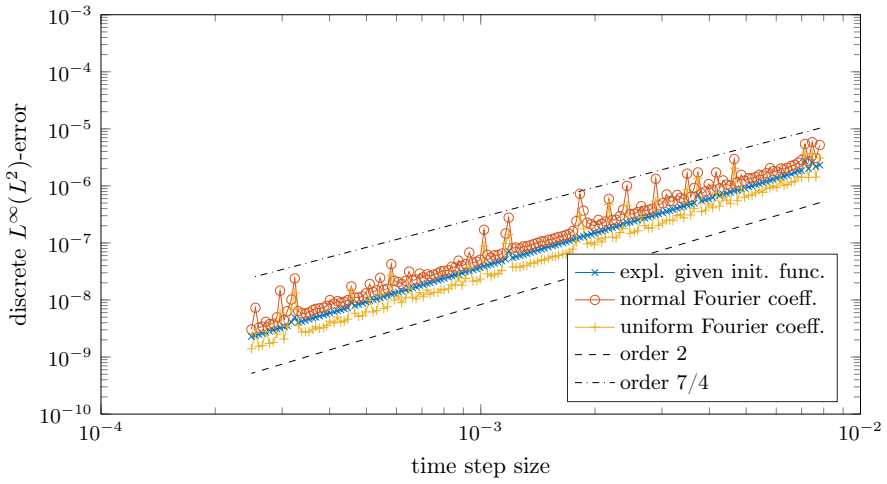


(c) $N = 4096$.

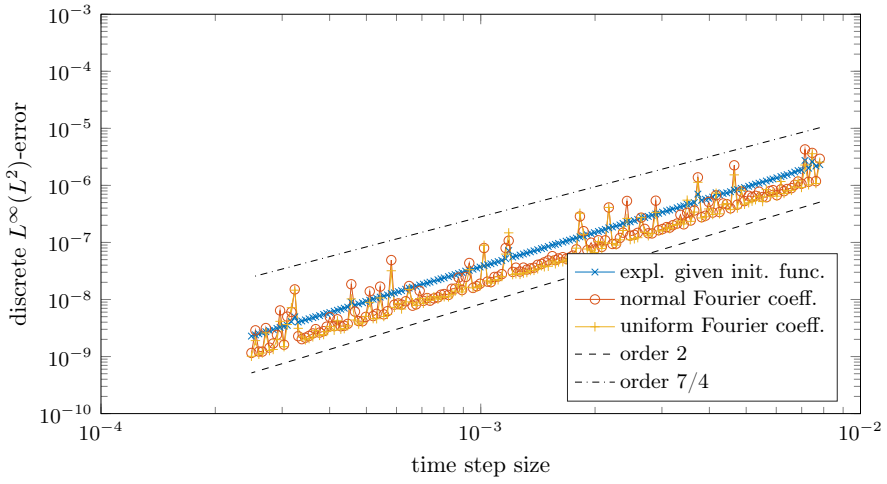
Figure 6.7.: Errors of the Strang splitting for initial functions being almost in H^4 in the focusing case for N space grid points.



(a) $N = 1024$.

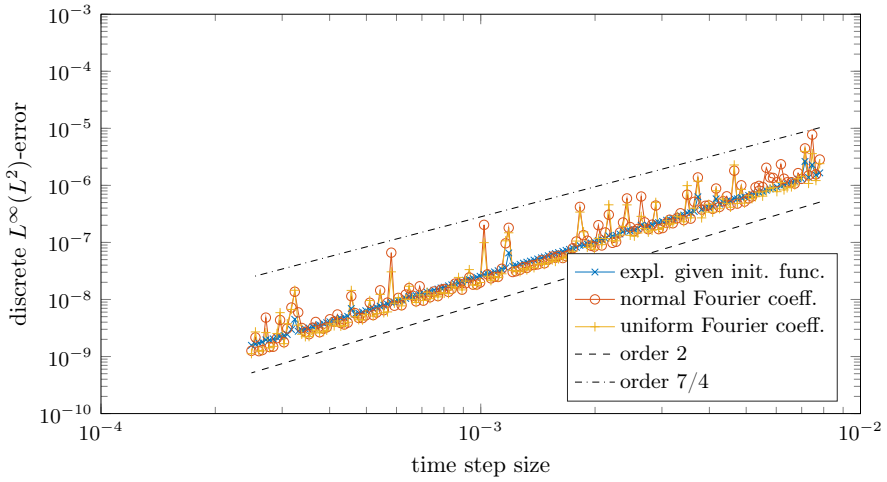


(b) $N = 2048$.

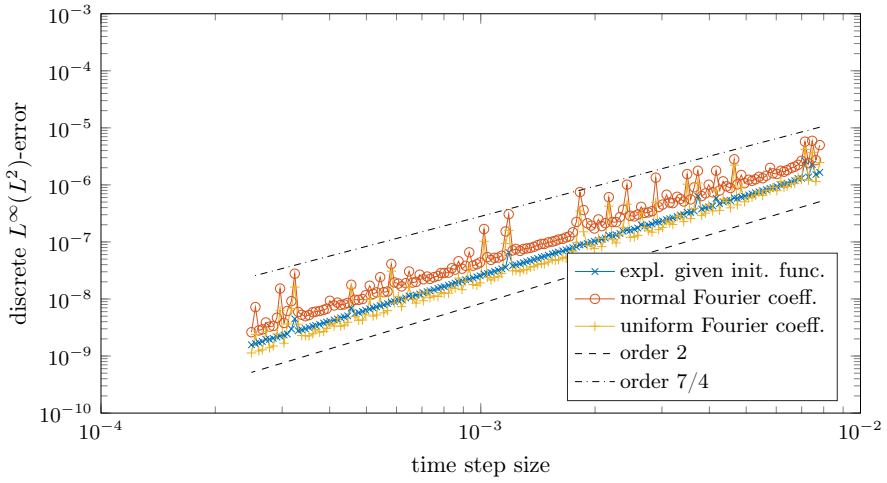


(c) $N = 4096$.

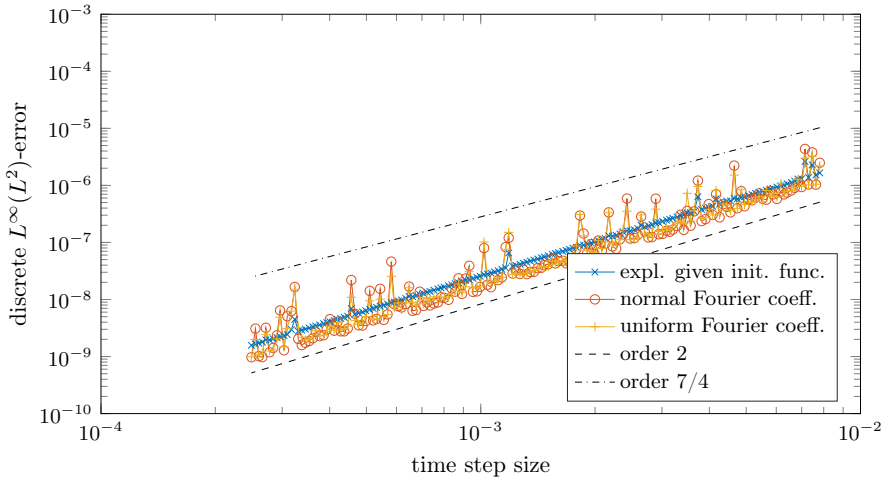
Figure 6.8.: Errors of the Strang splitting for initial functions being almost in $H^{7/2}$ in the defocusing case for N space grid points.



(a) $N = 1024$.

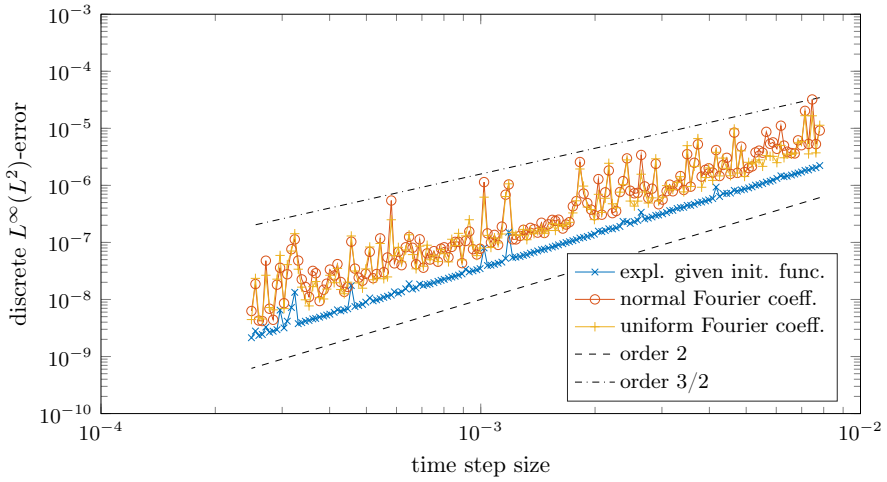


(b) $N = 2048$.

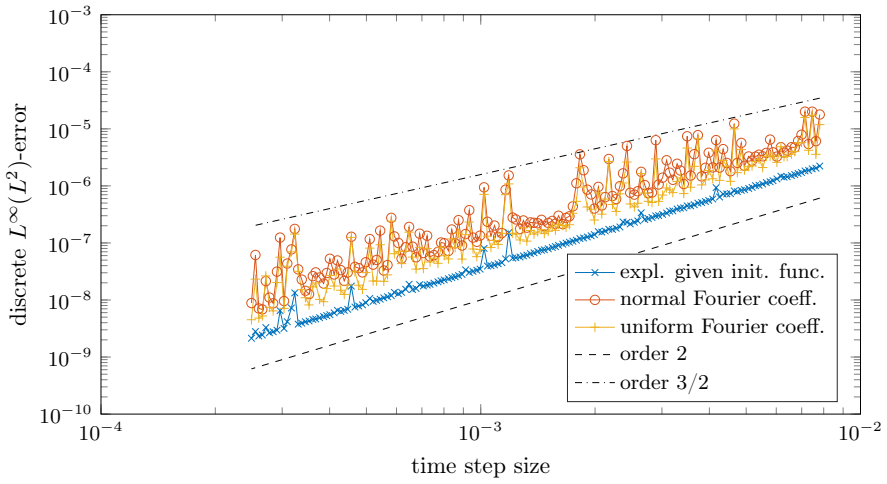


(c) $N = 4096$.

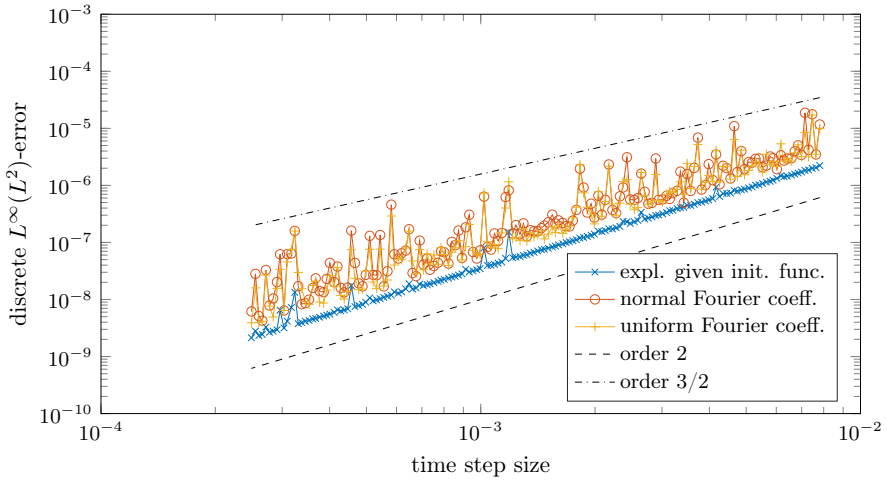
Figure 6.9.: Errors of the Strang splitting for initial functions being almost in $H^{7/2}$ in the focusing case for N space grid points.



(a) $N = 1024$.

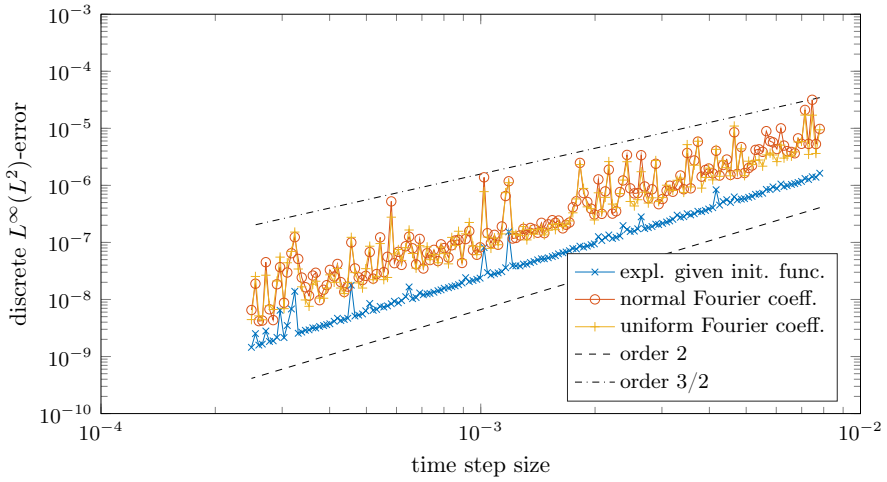


(b) $N = 2048$.

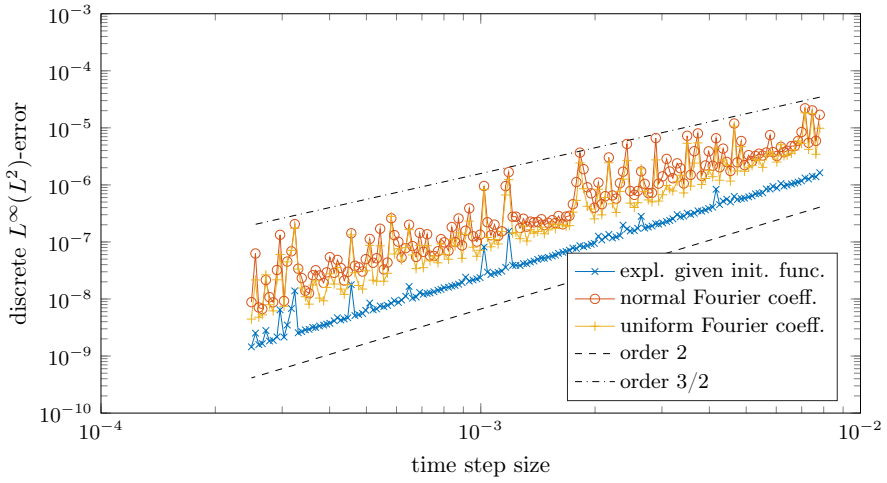


(c) $N = 4096$.

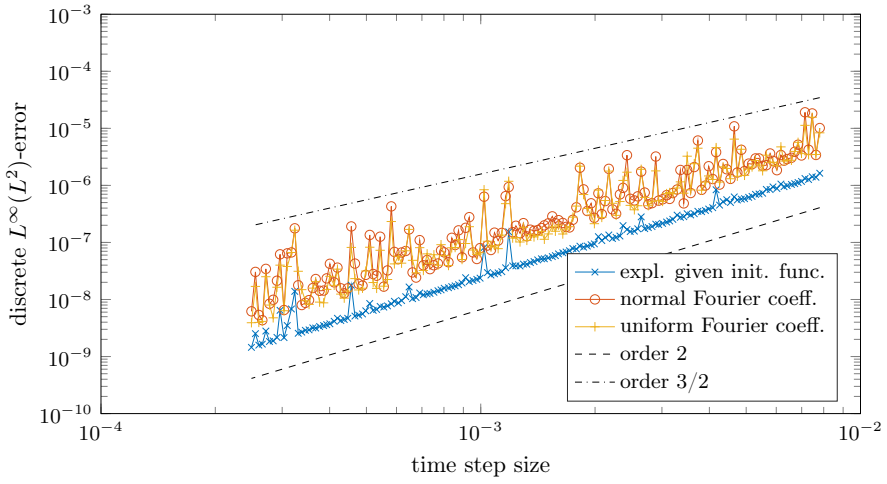
Figure 6.10.: Errors of the Strang splitting for initial functions being almost in H^3 in the defocusing case for N space grid points.



(a) $N = 1024$.

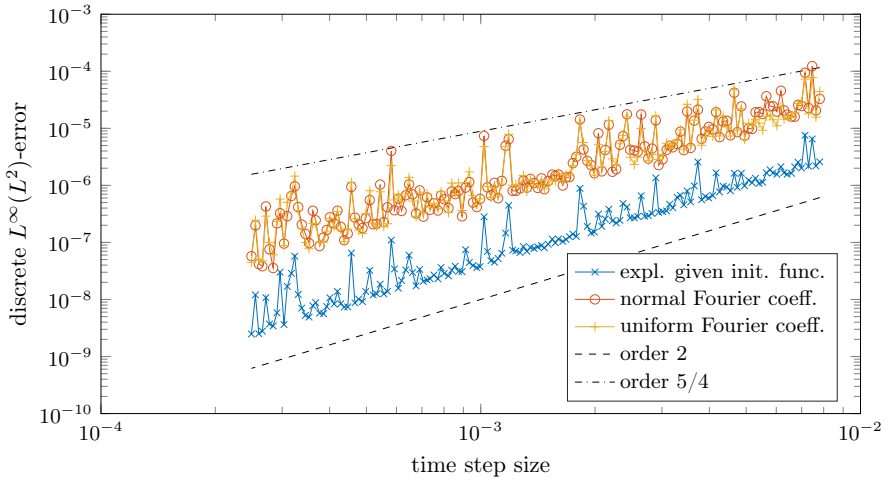


(b) $N = 2048$.

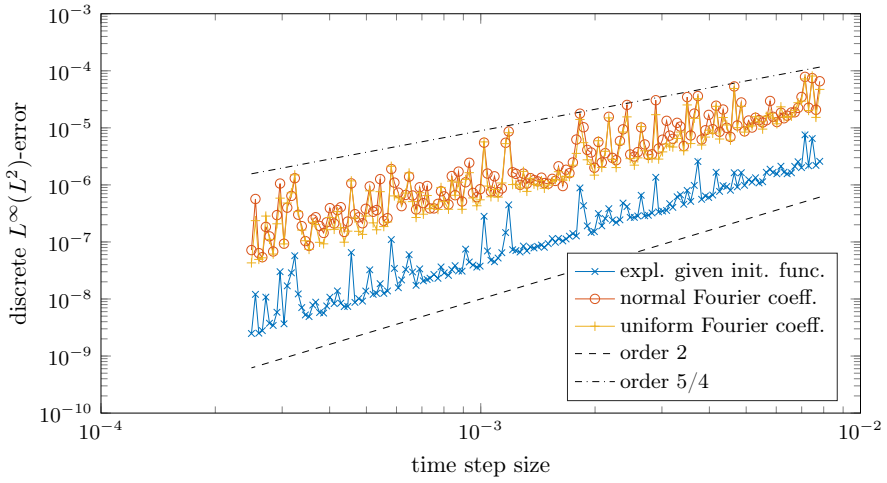


(c) $N = 4096$.

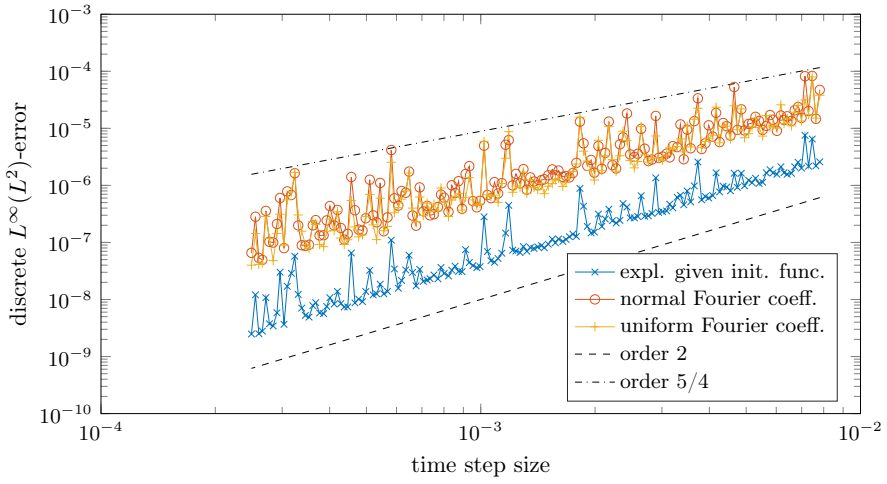
Figure 6.11.: Errors of the Strang splitting for initial functions being almost in H^3 in the focusing case for N space grid points.



(a) $N = 1024$.

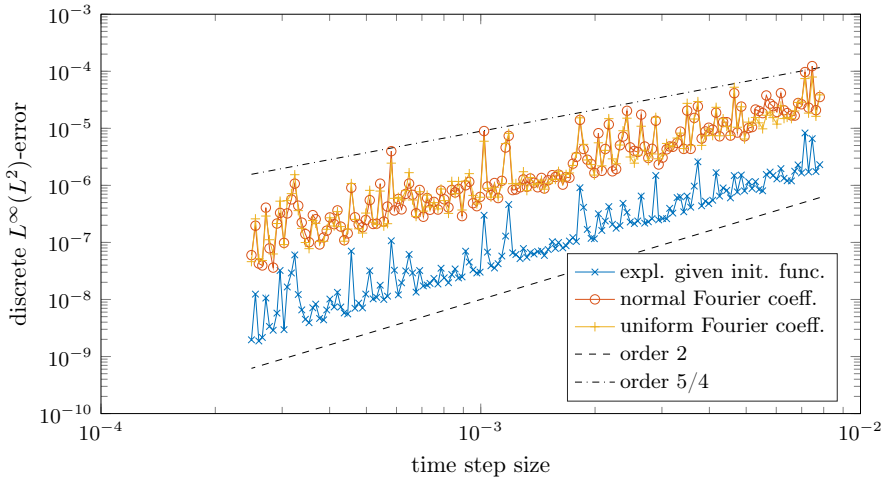


(b) $N = 2048$.

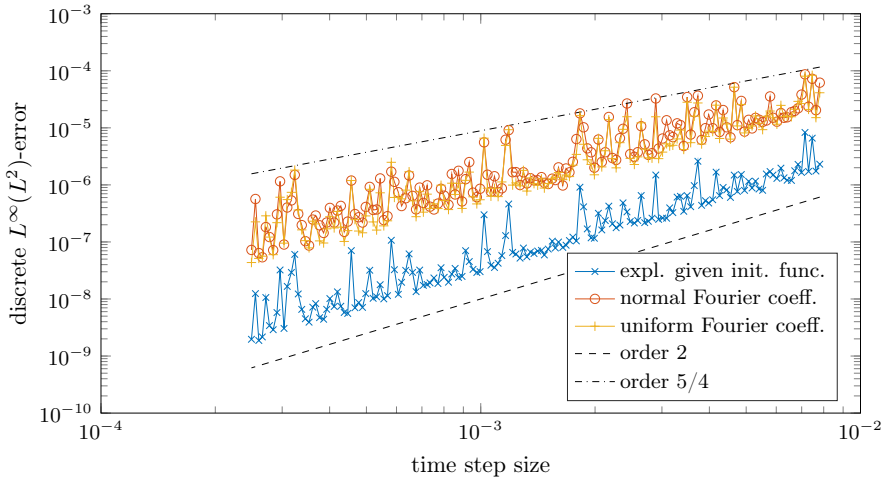


(c) $N = 4096$.

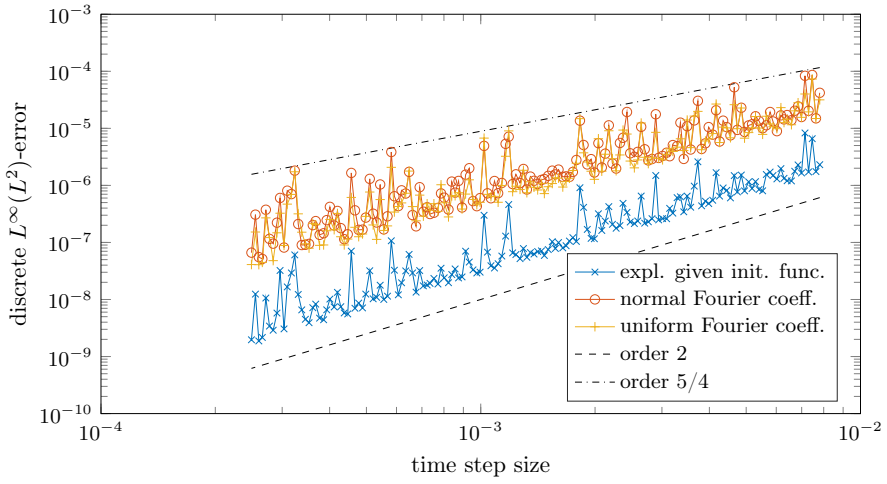
Figure 6.12.: Errors of the Strang splitting for initial functions being almost in $H^{5/2}$ in the defocusing case for N space grid points.



(a) $N = 1024$.

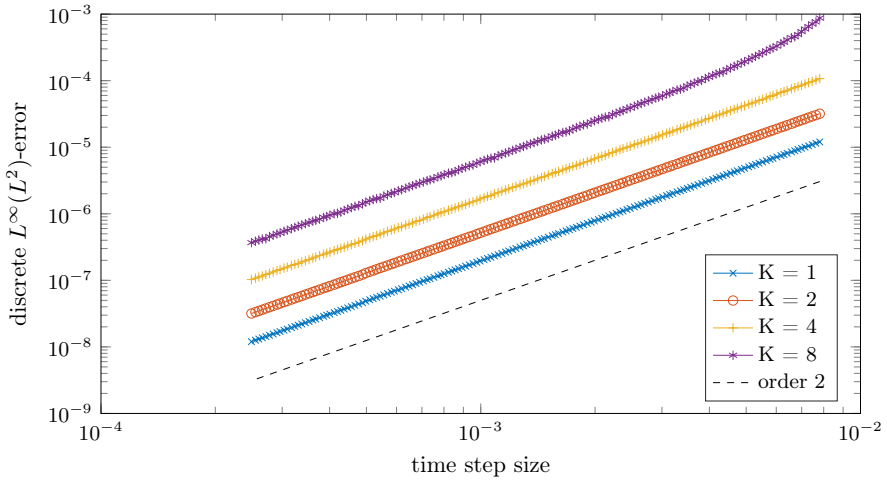


(b) $N = 2048$.

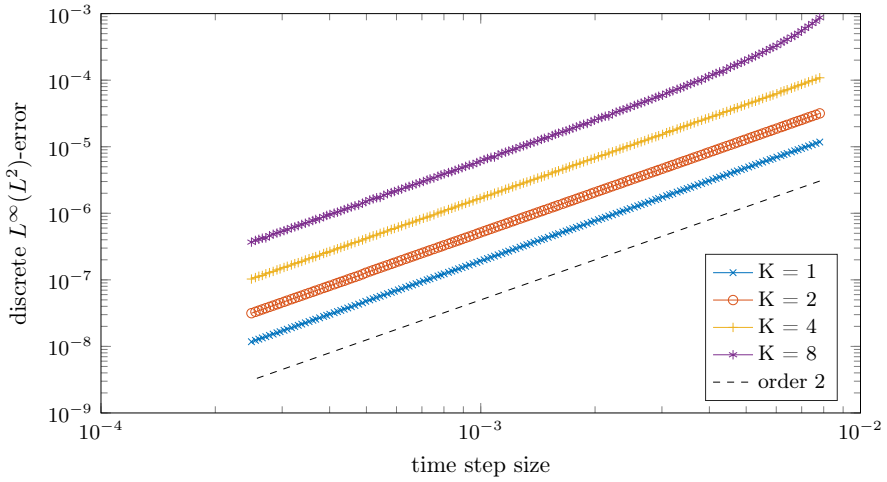


(c) $N = 4096$.

Figure 6.13.: Errors of the Strang splitting for initial functions being almost in $H^{5/2}$ in the focusing case for N space grid points.



(a) Defocusing case.



(b) Focusing case.

Figure 6.14.: Errors of the Strang splitting with smooth oscillating initial functions, using the oscillating factor K.

Part III.

An ADI splitting for the Maxwell equations

7. The Maxwell equations and their solutions

In this chapter we introduce the Maxwell equations and give a short overview over the properties of their solutions. In Section 7.1 we present the problem we deal with and the questions we tackle in this part of the thesis. The functional analytic background for our analysis and our splitting operators are described in Section 7.2. This section also includes the proofs of some embedding properties. Semigroup generation properties of the Maxwell operators and properties of the solutions we gain by using them are discussed in Section 7.3.

7.1. The Maxwell equations

On a spatial domain $Q \subseteq \mathbb{R}^3$ we consider for $t \geq 0$ the *Maxwell equations*

$$\partial_t \mathbf{E}(t) = \frac{1}{\varepsilon} \operatorname{curl} \mathbf{H}(t) - \frac{1}{\varepsilon}(\sigma \mathbf{E}(t) + \mathbf{J}_0(t)) \qquad \text{in } Q, \tag{7.1a}$$

$$\partial_t \mathbf{H}(t) = -\frac{1}{\mu} \operatorname{curl} \mathbf{E}(t) \qquad \text{in } Q, \tag{7.1b}$$

$$\operatorname{div}(\varepsilon \mathbf{E}(t)) = \rho(t) \qquad \text{in } Q, \tag{7.1c}$$

$$\operatorname{div}(\mu \mathbf{H}(t)) = 0 \qquad \text{in } Q, \tag{7.1d}$$

supplemented by the boundary conditions

$$\mathbf{E}(t) \times \nu = 0 \qquad \text{on } \partial Q, \tag{7.1e}$$

$$\mu \mathbf{H}(t) \cdot \nu = 0 \qquad \text{on } \partial Q, \tag{7.1f}$$

and the initial conditions

$$\mathbf{E}(0) = \mathbf{E}_0, \qquad \mathbf{H}(0) = \mathbf{H}_0 \qquad \text{in } Q, \tag{7.1g}$$

where ν is the outer unit normal vector. In the following, Q is the interior of a three-dimensional cuboid whose edges are parallel to the coordinate axes, which ensures the unique existence of ν in almost all boundary points. The unknowns $\mathbf{E}(t, x) \in \mathbb{R}^3$ and

$\mathbf{H}(t,x) \in \mathbb{R}^3$ are the electric and magnetic field, respectively. The electric permittivity and the magnetic permeability are denoted by $\varepsilon(x) \in (0,\infty)$ and $\mu(x) \in (0,\infty)$, respectively. Furthermore, $\mathbf{J}_0(t,x) \in \mathbb{R}^3$ is the external electric current density, $\sigma(x) \geq 0$ the electric conductivity and $\rho(t,x) \in \mathbb{R}$ the electric charge density. The initial fields $\mathbf{E}_0$ and $\mathbf{H}_0$ belong to $L^2(Q,\mathbb{R})^3$. We treat the case of perfectly conducting boundary conditions $\mathbf{E}(t) \times \nu = 0$ and $\mu\mathbf{H}(t) \cdot \nu = 0$. They describe the situation that the electric flux lines are on the boundary perpendicular to the surface and that the magnetic flux lines never cross the boundary.

Equation (7.1a) is Ampère's circuital law that relates the change of the electric field to the induced magnetic field, including an external current density and a damping term caused by electric conductivity. Faraday's law of induction (7.1b) connects the change of the magnetic field to the induced electric field. Gauss's law (7.1c) says that the electric flux that leaves a volume is proportional to the charge inside. Gauss's law for magnetism in (7.1d) states that there no magnetic charges exist and that the electric flux through every closed surface is zero.

Let $\tau > 0$. We set $t_n := n\tau$ for $n \in \mathbb{N}_0$. The *alternating direction implicit (ADI) splitting scheme* $S^I_{\tau,n+1}$ we investigate is given by

$$\begin{aligned} S^I_{\tau,n+1} w :&= (I - \tfrac{\tau}{2}B)^{-1}(I + \tfrac{\tau}{2}A)\cdot \\ &\quad \cdot \Big[(I - \tfrac{\tau}{2}A)^{-1}(I + \tfrac{\tau}{2}B)w - \tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1}), 0)\Big], \end{aligned}$$

see also Section 8.4. The conductivity σ is included in the splitting operators A and B that are defined by

$$A := \begin{pmatrix} -\frac{\sigma}{2\varepsilon} I & \frac{1}{\varepsilon} C_1 \\ \frac{1}{\mu} C_2 & 0 \end{pmatrix} \quad \text{and} \quad B := \begin{pmatrix} -\frac{\sigma}{2\varepsilon} I & -\frac{1}{\varepsilon} C_2 \\ -\frac{1}{\mu} C_1 & 0 \end{pmatrix},$$

$$C_1 := \begin{pmatrix} 0 & 0 & \partial_2 \\ \partial_3 & 0 & 0 \\ 0 & \partial_1 & 0 \end{pmatrix} \quad \text{and} \quad C_2 := \begin{pmatrix} 0 & \partial_3 & 0 \\ 0 & 0 & \partial_1 \\ \partial_2 & 0 & 0 \end{pmatrix}.$$

The sum of A and B is the Maxwell operator that governs (7.1).

This splitting scheme has been introduced for $\sigma = 0$ in [75]. It is efficient, stable and formally of second order. In [37] an error analysis in L^2 has been done for $\sigma = 0$, $\rho = 0$ and $\mathbf{J}_0 = 0$, where only zero divergence conditions have been considered. In the thesis at hand we treat in Theorem 9.3 the full problem with nontrivial charge densities, conductivity and external current densities. We thus have to include an inhomogeneity into the numerical scheme, see above. Furthermore, we add a convergence analysis in an H^{-1}-setting under weaker assumptions on the data, see Theorem 9.5. In both situations the result of the numerical scheme converges to the solution to (7.1) with order two in the time step size.

The solution to problem (7.1) fulfils the divergence conditions

$$\operatorname{div}(\varepsilon \mathbf{E}(t)) = \operatorname{div}(\varepsilon \mathbf{E}_0) - \int_0^t \operatorname{div}(\sigma \mathbf{E}(s) + \mathbf{J}_0(s))\,\mathrm{d}s, \tag{7.2a}$$

$$\operatorname{div}(\mu \mathbf{H}(t)) = \operatorname{div}(\mu \mathbf{H}_0) \tag{7.2b}$$

for $t \geq 0$, see Section 7.3. We show that the divergence of the numerical solution differs in L^2 and in H^{-1} only linearly in the time step size from (7.2), see Theorem 9.9 and 9.6. Again, the result in H^{-1} requires less regularity of the data.

Throughout, we assume that the material coefficients satisfy the *general assumptions*

$$\begin{aligned} \varepsilon, \mu, \sigma &\in W^{1,\infty}(Q, \mathbb{R}), \\ \varepsilon, \mu \geq \delta \quad &\text{for a constant } \delta > 0, \qquad \sigma \geq 0. \end{aligned}$$

Recall that the Sobolev space $W^{1,\infty}(Q, \mathbb{R})$ coincides with the space of bounded Lipschitz conditions functions. In some results we have to pose slightly stronger assumptions on the coefficients. The assumptions on the data $\mathbf{E}_0$, $\mathbf{H}_0$ and $\mathbf{J}_0$ differ from theorem to theorem. Roughly speaking we will assume at least that they belong to H^1 and satisfy the boundary conditions. We emphasize that ε, μ, σ, $\mathbf{J}_0$, $\mathbf{E}_0$ and $\mathbf{H}_0$ are given functions, while ρ is not given in advance, but will be determined by the solution to (7.1). Assumptions on the regularity of $\rho(t)$ for $t \geq 0$ are therefore constraints on the regularity of the solution.

Remark 7.1. *The speed of light is proportional to $\frac{1}{\sqrt{\varepsilon\mu}}$ with a constant depending on the units. An infimum of the electric permittivity or the magnetic permeability of zero would therefore lead to an arbitrarily large speed of light. Thus, it is physically reasonable that ε and μ are bounded away from zero.*

We describe the cuboid by

$$Q := (a_1^-, a_1^+) \times (a_2^-, a_2^+) \times (a_3^-, a_3^+)$$

with $a_j^\pm \in \mathbb{R}$ and $a_j^- < a_j^+$ for $j = 1, 2, 3$. We denote its boundary by $\Gamma := \partial Q$ and its outer unit normal vector by ν. We introduce the notations

$$\Gamma_j^\pm := \{(x_1, x_2, x_3) \in \Gamma \mid x_j = a_j^\pm\}$$

and $\Gamma_j := \Gamma_j^- \cup \Gamma_j^+$ for $j = 1, 2, 3$ and define

$$d_{\min} := \min_{j \in \{1,2,3\}} (a_j^+ - a_j^-).$$

We abbreviate $L^2(Q) := L^2(Q, \mathbb{R})$ and so on.

7.2. The functional analytic setting

7.2.1. Function spaces for the Maxwell operator

We first observe that the general assumptions on the coefficient functions ε and μ imply $\frac{1}{\varepsilon}, \frac{1}{\mu} \in W^{1,\infty}(Q)$ and that $\frac{1}{\varepsilon}$ and $\frac{1}{\mu}$ are bounded away from zero.

The following lemma states that some Sobolev spaces are invariant under multiplication with certain functions. We use this fact later on for the material coefficient functions.

Lemma 7.2. *(a) Let $\psi \in W^{1,\infty}(Q)$. Then the mapping $f \mapsto \psi f$ is continuous on $H^1(Q)$ and we have*

$$\|\psi f\|_{H^1} \leq c \, \|\psi\|_{W^{1,\infty}} \, \|f\|_{H^1}$$

for all $f \in H^1(Q)$.

(b) Let $\psi \in W^{1,\infty}(Q) \cap W^{2,3}(Q)$. Then the mapping $f \mapsto \psi f$ is continuous on $H^2(Q)$ and we have

$$\|\psi f\|_{H^2} \leq c \, \|\psi\|_{W^{1,\infty} \cap W^{2,3}} \, \|f\|_{H^2}$$

for all $f \in H^2(Q)$.

Proof:
(a) Young's inequality yields

$$\|\psi f\|_{H^1}^2 \leq c \Big(\|\psi f\|_{L^2}^2 + \sum_{k=1}^{3} \|f \partial_k \psi\|_{L^2}^2 + \sum_{k=1}^{3} \|\psi \partial_k f\|_{L^2}^2 \Big) \leq c \, \|f\|_{W^{1,\infty}}^2 \, \|f\|_{H^1}^2 \, .$$

(b) Using Hölder's inequality and the Sobolev embedding $H^1(Q) \hookrightarrow L^6(Q)$, we estimate

$$\|f \partial_{kl} \psi\|_{L^2} \leq \|f\|_{L^6} \, \|\partial_{kl} \psi\|_{L^3} \leq c \, \|f\|_{H^1} \, \|\psi\|_{W^{2,3}}$$

for all $k, l \in \{1, 2, 3\}$ and thus with Young's inequality

$$\begin{aligned} \|\psi f\|_{H^2}^2 \leq c \Big(& \|\psi f\|_{L^2}^2 + \sum_{k=1}^{3} \|f \partial_k \psi\|_{L^2}^2 + \sum_{k=1}^{3} \|\psi \partial_k f\|_{L^2}^2 \\ & + \sum_{k,l=1}^{3} \|f \partial_{kl} \psi\|_{L^2}^2 + \sum_{k,l=1}^{3} \|(\partial_k \psi)(\partial_l f)\|_{L^2}^2 + \sum_{k,l=1}^{3} \|\psi \partial_{kl} f\|_{L^2}^2 \Big) \\ & \leq c \big(\|\psi\|_{W^{1,\infty}}^2 + \|\psi\|_{W^{2,3}}^2 \big) \, \|f\|_{H^2}^2 \leq c \big(\|\psi\|_{W^{1,\infty}} + \|\psi\|_{W^{2,3}} \big)^2 \, \|f\|_{H^2}^2 \, , \end{aligned}$$

which is the claimed statement. □

We define the space $X := L^2(Q)^6$ and equip it with the weighted inner product

$$\big((u, v) \mid (\varphi, \psi) \big)_X := \int_Q \big(\varepsilon u \cdot \varphi + \mu v \cdot \psi \big) \, \mathrm{d}x$$

for $(u,v),(\varphi,\psi) \in X$, which induces a norm $\|\cdot\|_X$. Due to the general assumptions on ε and μ this so-called "energy norm" is equivalent to the L^2-norm. We introduce the Hilbert spaces

$$H(\operatorname{curl},Q) := \{u \in L^2(Q)^3 \mid \operatorname{curl} u \in L^2(Q)^3\}, \qquad \|u\|^2_{\operatorname{curl}} := \|u\|^2_{L^2} + \|\operatorname{curl} u\|^2_{L^2}$$

and

$$H(\operatorname{div},Q) := \{u \in L^2(Q)^3 \mid \operatorname{div} u \in L^2(Q)\}, \qquad \|u\|^2_{\operatorname{div}} := \|u\|^2_{L^2} + \|\operatorname{div} u\|^2_{L^2}\,.$$

We moreover define

$$H_0(\operatorname{curl},Q) := \overline{C_c^\infty(Q)}^{\|\cdot\|_{\operatorname{curl}}}.$$

In the next result we collect well-known facts about traces.

Proposition 7.3. *(a) The Dirichlet trace $u \mapsto u|_\Gamma$ on $C(\overline{Q})^3 \cap H^1(Q)^3$ has a unique continuous surjective extension $\operatorname{tr} : H^1(Q)^3 \to H^{1/2}(\Gamma)^3$ and the Neumann trace $u \mapsto \partial_\nu u$ is the continuous mapping $\operatorname{tr}_\nu : H^2(Q)^3 \to H^{1/2}(\Gamma)^3$.*

(b) The tangential trace $u \mapsto (u \times \nu)|_\Gamma$ on $C(\overline{Q})^3 \cap H^1(Q)^3$ has a unique continuous extension $\operatorname{tr}_t : H(\operatorname{curl},Q) \to H^{-1/2}(\Gamma)^3$. For all $u \in H(\operatorname{curl},Q)$ and $v \in H^1(Q)^3$ it holds

$$\int_Q \operatorname{curl} u \cdot v \,\mathrm{d}x = \int_Q u \cdot \operatorname{curl} v \,\mathrm{d}x - \langle \operatorname{tr}_t(u), v\rangle_{H^{-1/2}(Q)^3 \times H^{1/2}(Q)^3}\,.$$

(c) The normal trace $u \mapsto (u\cdot\nu)|_\Gamma$ on $C(\overline{Q})^3 \cap H^1(Q)^3$ has a unique continuous extension $\operatorname{tr}_n : H(\operatorname{div},Q) \to H^{-1/2}(\Gamma)$. For all $u \in H(\operatorname{div},Q)$ and $v \in H^1(Q)$ it holds

$$\int_Q \operatorname{div}(u) v \,\mathrm{d}x = \int_Q u \cdot \nabla v \,\mathrm{d}x + \langle \operatorname{tr}_n(u), v\rangle_{H^{-1/2}(Q)\times H^{1/2}(Q)}\,.$$

(d) The space

$$C_c^\infty(\overline{Q}) = \left\{ f|_Q \;\middle|\; f \in C_c^\infty(\mathbb{R}^3)\right\}$$

is dense in $H^1(Q)$, $H(\operatorname{div},Q)$ and $H(\operatorname{curl},Q)$.

(e) Defining for $A \subseteq \Gamma$ the restricted trace $\operatorname{tr}_A(u) := \mathbb{1}_A \operatorname{tr}(u)$, we have

$$\begin{aligned} H_0(\operatorname{curl},Q) :&= \{u \in H(\operatorname{curl},Q) \mid \operatorname{tr}_t(u) = 0 \text{ on } \Gamma\} \\ &= \{u \in H(\operatorname{curl},Q) \mid \operatorname{tr}_{\Gamma_1}(u_2) = \operatorname{tr}_{\Gamma_1}(u_3) = \operatorname{tr}_{\Gamma_2}(u_1) \\ &\qquad\qquad = \operatorname{tr}_{\Gamma_2}(u_3) = \operatorname{tr}_{\Gamma_3}(u_1) = \operatorname{tr}_{\Gamma_3}(u_2) = 0\}. \end{aligned}$$

Proof:
The statement on the Dirichlet trace follows from the Sections 2.4 and 2.5 in [58]. The claims on the Neumann trace then follow by taking the derivatives. The parts (b), (c) and (d) can be found in Section IX.A.1.2 in [16]. The formulas in part (b) and (c) for the partial integration are seen with Green's formula. Part (e) can be seen by an approximation argument. □

To ease the notation we write in the following $u_1 = 0$ on Γ_2 for the property $\mathrm{tr}_{\Gamma_2}(u_1) = 0$, and so on. Furthermore, for $\widetilde{\Gamma} \subseteq \Gamma$ being a union a some of the faces of Q we set

$$H^1_{\widetilde{\Gamma}}(Q) := \{u \in H^1(Q) \mid \mathrm{tr}(u) = 0 \text{ on } \widetilde{\Gamma}\}.$$

It is clear that $H^1(Q)^3$ embeds continuously into $H(\mathrm{curl}, Q)$ and $H(\mathrm{div}, Q)$. The following proposition states the converse implication under an additional assumption, see Theorem 2.17 in [3].

Proposition 7.4. *Let $f \in H(\mathrm{div}, Q) \cap H(\mathrm{curl}, Q)$ and let either $\mathrm{tr}_t(f) = 0$ or $\mathrm{tr}_n(f) = 0$ on Γ. Then $f \in H^1(Q)^3$ and*

$$\|f\|_{H^1} \leq c\big(\|f\|_{L^2} + \|\mathrm{div}\, f\|_{L^2} + \|\mathrm{curl}\, f\|_{L^2}\big).$$

We note that for sufficiently regular functions the trace is multiplicative.

Lemma 7.5. *Let $p, q \in (1, \infty]$ with $\frac{1}{p} + \frac{1}{q} < 1$.*

(a) Let $f \in W^{1,p}(Q)$ and $g \in W^{1,q}(Q)$. Then we have

$$\mathrm{tr}(fg) = \mathrm{tr}(f)\,\mathrm{tr}(g).$$

(b) Let $f \in W^{1,p}(Q)$ and $h \in W^{1,q}(Q)^3$. Then we have

$$\mathrm{tr}(fh) = \mathrm{tr}(f)\,\mathrm{tr}(h), \qquad \mathrm{tr}_t(fh) = \mathrm{tr}(f)\,\mathrm{tr}_t(h) \qquad \text{and} \qquad \mathrm{tr}_n(fh) = \mathrm{tr}(f)\,\mathrm{tr}_n(h).$$

Proof:
To show part (a), we approximate f in $W^{1,p}(Q)$ and g in $W^{1,q}(Q)$ by functions f_n and g_n in $W^{1,\infty}(Q)$. We omit the respective approximation if $p = \infty$ or $q = \infty$. For f_n and g_n the result is true and it extends to f and g by the continuity of the trace operator since $fg \in W^{1,r}(Q)$ for $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. The statement of part (b) follows in the same way. □

We define the *Maxwell operator*

$$M := \begin{pmatrix} -\frac{\sigma}{\varepsilon} I & \frac{1}{\varepsilon}\,\mathrm{curl} \\ -\frac{1}{\mu}\,\mathrm{curl} & 0 \end{pmatrix} \tag{7.3}$$

with domain $D(M) := H_0(\mathrm{curl}, Q) \times H(\mathrm{curl}, Q)$: Observe that the electric boundary condition is included in this domain. We abbreviate

$$\begin{pmatrix} \big(M(\mathbf{E}, \mathbf{H})\big)_1 \\ \big(M(\mathbf{E}, \mathbf{H})\big)_2 \end{pmatrix} := M(\mathbf{E}, \mathbf{H}) = \begin{pmatrix} -\frac{\sigma}{\varepsilon}\mathbf{E} + \frac{1}{\varepsilon}\,\mathrm{curl}\,\mathbf{H} \\ -\frac{1}{\mu}\,\mathrm{curl}\,\mathbf{E} \end{pmatrix}$$

for $(\mathbf{E}, \mathbf{H}) \in D(M)$. We define, as usual

$$D(M^2) := \{(\mathbf{E}, \mathbf{H}) \in D(M) \mid M(\mathbf{E}, \mathbf{H}) \in D(M)\},$$

and so on. The above domain only contains the electric boundary conditions. The magnetic ones and the divergence conditions are encoded in the subspace

$$X_0 := \{(u,v) \in X \mid \operatorname{div}(\varepsilon u) = \operatorname{div}(\mu v) = 0, \ \ \operatorname{tr}_n(\mu v) = 0 \text{ on } \Gamma\}. \tag{7.4}$$

Here, the constraints are meant in the sense that the equations in Q hold true in $H^{-1}(Q)$, while the trace is zero in $H^{-1/2}(\Gamma)$, compare Proposition 7.3.

Lemma 7.6. *The subspace X_0 equipped with the norm $\|\cdot\|_X$ is a closed subspace of X.*

Proof:
Let $(u,v) \in X$. Since ε belongs to $W^{1,\infty}(Q)$ and

$$\operatorname{div}(\varepsilon u) = \nabla\varepsilon \cdot u + \varepsilon \operatorname{div}(u) \quad \Longleftrightarrow \quad \operatorname{div}(u) = \tfrac{1}{\varepsilon}\operatorname{div}(\varepsilon u) - \tfrac{1}{\varepsilon}\nabla\varepsilon \cdot u, \tag{7.5}$$

we see that $\operatorname{div}(\varepsilon u) \in L^2(Q)$ if and only if $\operatorname{div}(u) \in L^2(Q)$, and analogously for $\operatorname{div}(\mu v)$. This shows

$$X_0 \subseteq H(\operatorname{div}, Q) \times H(\operatorname{div}, Q) \subseteq X.$$

The closedness of X_0 in L^2 then follows from the closedness of the divergence and the continuity of the normal trace. □

If the charge density ρ is not zero, we need different spaces in view of (7.1c). We first introduce the space $H^1_{00}(Q)$ of all functions in $H^1(Q)$ such that for all faces $\widehat{\Gamma}$ of Q the Dirichlet traces on $\widehat{\Gamma}$ are contained in $H_0^{1/2}(\widehat{\Gamma})$. This means that the boundary values are zero on the edges of Q in a generalised sense. We need this property in some later proofs as a compatibility condition. Here, for a face $\widehat{\Gamma}$ of Q the space $H_0^{1/2}(\widehat{\Gamma})$ is the real interpolation space

$$H_0^{1/2}(\widehat{\Gamma}) := \big(L^2(\widehat{\Gamma}), H_0^1(\widehat{\Gamma})\big)_{1/2,2}.$$

Interpolation of the inclusion maps $L^2(\widehat{\Gamma}) \to L^2(\widehat{\Gamma})$ and $H_0^1(\widehat{\Gamma}) \to H^1(\widehat{\Gamma})$ yields the embedding

$$H_0^{1/2}(\widehat{\Gamma}) \hookrightarrow H^{1/2}(\widehat{\Gamma}).$$

We write $H_0^{1/2}(\widetilde{\Gamma})$ in the case that $\widetilde{\Gamma}$ is the union of some faces of Q and mean by the notation $u \in H_0^{1/2}(\widetilde{\Gamma})$ that u belongs to $u \in H_0^{1/2}(\widehat{\Gamma})$ for all $\widehat{\Gamma} \subseteq \widetilde{\Gamma}$. We then define the subspaces

$$X^{(0)}_{\operatorname{div}} := \{(u,v) \in X \mid \operatorname{div}(\mu v) = 0, \ \ \operatorname{tr}_n(\mu v) = 0 \text{ on } \Gamma, \ \ \operatorname{div}(\varepsilon u) \in L^2(Q)\} \tag{7.6a}$$

and

$$X^{(2)}_{\operatorname{div}} := \{(u,v) \in D(M^2) \mid \operatorname{div}(\mu v) = 0, \ \ \operatorname{tr}_n(\mu v) = 0 \text{ on } \Gamma, \ \ \operatorname{div}(\varepsilon u) \in H^1_{00}(Q)\} \tag{7.6b}$$

with the norms given by

$$\|(u,v)\|^2_{X^{(0)}_{\mathrm{div}}} := \|(u,v)\|^2_X + \|\mathrm{div}(\varepsilon u)\|^2_{L^2}$$

for $(u,v) \in X^{(0)}_{\mathrm{div}}$ and

$$\|(u,v)\|^2_{X^{(2)}_{\mathrm{div}}} := \|(u,v)\|^2_{D(M^2)} + \|\mathrm{div}(\varepsilon u)\|^2_{H^1} + \sum_{\widehat{\Gamma} \text{ face of } Q} \|\mathrm{div}(\varepsilon u)\|^2_{H_0^{1/2}(\widehat{\Gamma})}$$

for $(u,v) \in X^{(2)}_{\mathrm{div}}$. It is clear that $X^{(2)}_{\mathrm{div}}$ is continuously embedded into $X^{(0)}_{\mathrm{div}}$. We will use the spaces $X^{(0)}_{\mathrm{div}}$ and $X^{(2)}_{\mathrm{div}}$ depending on the regularity of ρ and thus on the contraints on the regularity of the solution to (7.1).

Lemma 7.7. *The spaces $(X^{(0)}_{\mathrm{div}}, \|\cdot\|_{X^{(0)}_{\mathrm{div}}})$ and $(X^{(2)}_{\mathrm{div}}, \|\cdot\|_{X^{(2)}_{\mathrm{div}}})$ are Hilbert spaces, and X_0 is a closed subspace of them. Moreover, $X^{(0)}_{\mathrm{div}}$ is embedded in $H(\mathrm{div}, Q)^2$, where the constant depends only on $\|\varepsilon\|_{W^{1,\infty}}$, $\|\mu\|_{W^{1,\infty}}$ and δ.*

Proof:
Clearly, the norms of $X^{(0)}_{\mathrm{div}}$ and $X^{(2)}_{\mathrm{div}}$ are given by an inner product. The norm $\|\cdot\|_{X^{(0)}_{\mathrm{div}}}$ is equivalent to the norm given by $\|u\|^2_{\mathrm{div}} + \|v\|^2_{\mathrm{div}}$ due to (7.5). Furthermore, the maps $v \mapsto \mathrm{div}(\mu v)$ and $v \mapsto \mathrm{tr}_n(\mu v)$ are continuous on $H(\mathrm{div}, Q)$. Therefore, the space $(X^{(0)}_{\mathrm{div}}, \|\cdot\|_{X^{(0)}_{\mathrm{div}}})$ is complete as it is isomorphic to a closed subspace of the Hilbert space $H(\mathrm{div}, Q)^2$. Further, let $(u_n, v_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $X^{(2)}_{\mathrm{div}}$. Since M is closed, (u_n, v_n) then has a limit (u,v) in $D(M)$. Moreover, $\mathrm{div}(\varepsilon u_n)$ converges to a function $\varphi \in H^1(Q)^3$ and to $\mathrm{div}(\varepsilon u)$ in $H^{-1}(Q)^3$, so that $\varphi = \mathrm{div}(\varepsilon u)$. Similarly, the traces of $\mathrm{div}(\varepsilon u_n)$ on each face $\widehat{\Gamma}$ of Q tend to a function ψ in $H_0^{1/2}(\widehat{\Gamma})$ and also to $\mathrm{tr}\big(\mathrm{div}(\varepsilon u)\big)$ in $H^{1/2}(\widehat{\Gamma})$, i.e. $\mathrm{tr}\big(\mathrm{div}(\varepsilon u)\big) = \psi$ in $\widehat{\Gamma}$. As for $X^{(0)}_{\mathrm{div}}$ one checks the magnetic conditions. The closedness of X_0 in $X^{(0)}_{\mathrm{div}}$ and in $X^{(2)}_{\mathrm{div}}$ follows from the continuity of $u \mapsto \mathrm{div}(\varepsilon u)$ from $H(\mathrm{div}, Q)$ to $L^2(Q)^3$. □

We use these spaces to define the *Maxwell operators*

$$M_0 : D(M_0) := D(M) \cap X_0 \to X, \tag{7.7a}$$

$$M^{(0)}_{\mathrm{div}} : D(M^{(0)}_{\mathrm{div}}) := D(M) \cap X^{(0)}_{\mathrm{div}} \to X, \tag{7.7b}$$

$$M^{(2)}_{\mathrm{div}} : D(M^{(2)}_{\mathrm{div}}) := D(M^3) \cap X^{(2)}_{\mathrm{div}} \to X, \tag{7.7c}$$

which are restrictions of M. We note that these operators and M differ by the respective electric divergence and the magnetic conditions. Furthermore, $M^{(2)}_{\mathrm{div}}$ incorporates two more degrees of regularity. This is necessary in Section 7.3 to show that the semigroup generated by M leaves $X^{(2)}_{\mathrm{div}}$ invariant. We define, analogously to above,

$$D(M_0^2) := \{(u,v) \in D(M_0) \mid M(u,v) \in D(M_0)\},$$

$$D\big((M_{\mathrm{div}}^{(0)})^2\big) := \{(u,v) \in D(M_{\mathrm{div}}^{(0)}) \mid M(u,v) \in D(M_{\mathrm{div}}^{(0)})\}.$$

Our next goal is to show embedding properties of the domains of the Maxwell operators. For this we prove two auxiliary lemmas. The first one allows us to take a limit of boundary integrals.

Lemma 7.8. *We define for all $\kappa \in (0, d_{\min}/2)$ the cuboid*

$$Q_\kappa := \{(x_1,x_2,x_3) \in Q \mid \operatorname{dist}\big((x_1,x_2,x_3),\Gamma\big) > \kappa\}.$$

Then we have for $f \in H^1(Q)$ that

$$\lim_{\kappa\to 0} \int_{\partial Q_\kappa} |f|^2 \,\mathrm{d}\sigma = \int_\Gamma |f|^2 \,\mathrm{d}\sigma.$$

Proof:
Let $f \in H^1(Q)$. We define for all $\kappa \in (0, d_{\min}/2)$

$$I_\kappa(f) := \int_{\partial Q_\kappa} |f|^2 \,\mathrm{d}\sigma \qquad \text{and} \qquad I(f) := \int_\Gamma |f|^2 \,\mathrm{d}\sigma.$$

Let $(f_n)_{n\in\mathbb{N}}$ in $C^1(\overline{Q})$ be such that $f_n \to f$ in $H^1(Q)$ as $n \to \infty$. Let $\eta > 0$. With the continuity of the trace from $H^1(Q_\kappa)$ to $H^{1/2}(\partial Q_\kappa)$ we deduce

$$\begin{aligned}
|I_\kappa(f_n) - I_\kappa(f)| &\le \int_{\partial Q_\kappa} |f_n|\,|f_n - f| \,\mathrm{d}\sigma + \int_{\partial Q_\kappa} |f_n - f|\,|f| \,\mathrm{d}\sigma \\
&\le \big(\|f_n\|_{L^2(\partial Q_\kappa)} + \|f\|_{L^2(\partial Q_\kappa)}\big)\,\|f_n - f\|_{L^2(\partial Q_\kappa)} \\
&\le c\big(\|f_n\|_{H^1(Q_\kappa)} + \|f\|_{H^1(Q_\kappa)}\big)\,\|f_n - f\|_{H^1(Q_\kappa)} \\
&\le c\big(\|f_n\|_{H^1(Q)} + \|f\|_{H^1(Q)}\big)\,\|f_n - f\|_{H^1(Q)} \\
&\longrightarrow 0
\end{aligned}$$

as $n \to \infty$. Thus, we can choose an index $n_1 = n_1(\eta) \in \mathbb{N}$ independent of κ such that $|I_\kappa(f_n) - I_\kappa(f)| \le \eta$. In the same way we can choose an $n_2 = n_2(\eta) \in \mathbb{N}$ such that $|I(f_n) - I(f)| \le \eta$. From now on let $n \ge \max\{n_1, n_2\}$ be fixed. Analogously to $\Gamma_j^\pm$, $j = 1,2,3$, we define

$$\begin{aligned}
\partial Q_{\kappa,j}^- &:= \{(x_1,x_2,x_3) \in \partial Q_\kappa \mid x_j = a_j^- + \kappa\}, \\
\partial Q_{\kappa,j}^+ &:= \{(x_1,x_2,x_3) \in \partial Q_\kappa \mid x_j = a_j^+ - \kappa\}.
\end{aligned}$$

We define the set

$$\Gamma_{3,\kappa}^- := [a_1^- + \kappa, a_1^+ - \kappa] \times [a_2^- + \kappa, a_2^+ - \kappa] \times \{a_3^-\}$$

and so on. From

$$\begin{aligned} S_{1,n} &= \int_{\Gamma_{3,\kappa}^-} \left(\left| f_n(t, a_3^-) \right|^2 - \left| f_n(t + a_3^- + \kappa) \right|^2 \right) \mathrm{d}t \\ &\leq 2 \left\| f_n \right\|_{L^\infty} \int_{\Gamma_{3,\kappa}^-} \left| f_n(t, a_3^-) - f_n(t, a_3^- + \kappa) \right| \mathrm{d}t \\ &\leq 2 \left\| f_n \right\|_{L^\infty} \left\| f_n' \right\|_{L^\infty} \kappa (a_1^+ - a_1^-)(a_2^+ - a_2^-) \end{aligned}$$

and so on we infer

$$|I_\kappa(f_n) - I(f_n)| \leq c\kappa \left\| f_n \right\|_{L^\infty}^2 + \sum_{j=1}^{6} S_{j,n} \leq c_n \kappa \longrightarrow 0$$

as $\kappa \to 0$ for n fixed. The estimate

$$|I_\kappa(f) - I(f)| \leq |I_\kappa(f) - I_\kappa(f_n)| + |I_\kappa(f_n) - I(f_n)| + |I(f_n) - I(f)| \leq 3\eta$$

for κ small enough finishes the proof. □

We continue with a lemma on the regularity of the solutions of two integral equations.

Lemma 7.9. *Let $f \in L^2(Q)$.*

(a) Let $\widetilde{\Gamma}$ be the union of one or two of the sets Γ_1, Γ_2 and Γ_3, and $\widetilde{\Gamma}' = \Gamma \setminus \widetilde{\Gamma}$. Furthermore, let

$$D_0 := \{ u \in H^2(Q) \cap H^1_{\widetilde{\Gamma}}(Q) \mid \partial_\nu u = 0 \text{ on } \widetilde{\Gamma}' \}.$$

Then there exists a unique function $u \in H^1_{\widetilde{\Gamma}}(Q)$ such that

$$\int_Q u\varphi \,\mathrm{d}x + \int_Q \nabla u \cdot \nabla \varphi \,\mathrm{d}x = \int_Q f\varphi \,\mathrm{d}x$$

for all $\varphi \in H^1_{\widetilde{\Gamma}}(Q)$. Additionally, we have $u \in D_0$ and $u - \Delta u = f$. Finally, the H^2-norm and the graph norm of Δ are equivalent on D_0.

(b) Let $\widetilde{\Gamma}$ be the union of exact two of the sets Γ_1, Γ_2 and Γ_3, and $\widetilde{\Gamma}' = \Gamma \setminus \widetilde{\Gamma}$. Furthermore, let

$$D := H^2(Q) \cap H^1_{\widetilde{\Gamma}}(Q)$$

and $g \in L^2(\widetilde{\Gamma}')$. Then there exists a unique function $v \in H^1_{\widetilde{\Gamma}}(Q)$ such that

$$\int_Q v\varphi \,\mathrm{d}x + \int_Q \nabla v \cdot \nabla \varphi \,\mathrm{d}x = \int_Q f\varphi \,\mathrm{d}x + \int_{\widetilde{\Gamma}'} g\varphi \,\mathrm{d}\sigma \tag{7.8}$$

for all $\varphi \in H^1_{\widetilde{\Gamma}}(Q)$. If $g \in H_0^{1/2}(\widetilde{\Gamma}')$, then we additionally have $v \in D$, $v - \Delta v = f$, $\partial_\nu v = g$ on $\widetilde{\Gamma}'$ and

$$\|v\|_{H^2} \leq c\big(\|f\|_{L^2} + \|g\|_{H_0^{1/2}(\widetilde{\Gamma}')} \big).$$

PROOF:
We only show part (b) since (a) was shown in Lemma 3.6 in [37].

1) First we show that problem (7.8) has a unique solution in $H^1_{\widetilde{\Gamma}}(Q)$. We define the bilinear form $B : H^1_{\widetilde{\Gamma}}(Q) \times H^1_{\widetilde{\Gamma}}(Q) \to \mathbb{R}$ and the linear functional $F : H^1_{\widetilde{\Gamma}}(Q) \to \mathbb{R}$ by

$$B(u,v) := \int_Q uv \,\mathrm{d}x + \int_Q \nabla u \cdot \nabla v \,\mathrm{d}x \qquad \text{and}$$
$$F(u) := \int_Q fu \,\mathrm{d}x + \int_\Gamma gu \,\mathrm{d}\sigma.$$

For all $u, v \in H^1_{\widetilde{\Gamma}}(Q)$ we obtain the relations

$$|B(u,v)| \le 2 \|u\|_{H^1} \|v\|_{H^1}, \qquad B(u,u) = \|u\|^2_{H^1},$$
$$|F(u)| \le \|f\|_{L^2} \|u\|_{L^2} + \|g\|_{L^2(\widetilde{\Gamma}')} \|u\|_{L^2(\Gamma)} \le c\big(\|f\|_{L^2} + \|g\|_{L^2(\widetilde{\Gamma}')}\big) \|u\|_{H^1},$$

using $H^1(Q) \hookrightarrow L^2(\Gamma)$ for the last estimate. The Lemma of Lax-Milgram then yields a unique solution $\widetilde{v}$ in $H^1_{\widetilde{\Gamma}}(Q)$ to (7.8).

2) Next we prove that for $g \in H_0^{1/2}(\widetilde{\Gamma}')$ there exists a function $w \in H^2(Q)$ with $\partial_\nu w = g$ on $\widetilde{\Gamma}'$ and $w = 0$ on $\widetilde{\Gamma}$. Let without loss of generality $\widetilde{\Gamma}' = \Gamma_1$. Let $R \subseteq \mathbb{R}^2$ be a rectangle that is congruent to one of the two congruent parts of $\widetilde{\Gamma}'$ and let Δ_R be the Dirichlet Laplacian on R with domain $D(\Delta_R) = H^2(R) \cap H_0^1(R)$. Without further mentioning we use Δ_R on Γ_1^- and on Γ_1^+, i.e. with $R = \Gamma_1^-$ and $R = \Gamma_1^+$.

It is well-known that the spectrum of Δ_R consists only of finitely many discrete eigenvalues on the negative real axis without zero, compare e.g. Lemma 6.2.1 in [17] for the situation of a cube. Therefore, Δ_R is invertible. Furthermore, Δ_R is self-adjoint since it is symmetric and has its spectrum on the real axis. So, we can define $(-\Delta_R)^{1/2}$ and $(-\Delta_R)^{-1/2}$ with the functional calculus for self-adjoint operators and these operators again have a discrete spectrum and are self-adjoint, see Theorem VII.1 in [64]. Hence, $(-\Delta_R)^{1/4}$ can be defined in the same way. From

$$\big((-\Delta_R)^{1/2} h, h\big)_{L^2} = \big((-\Delta_R)^{1/4} h, (-\Delta_R)^{1/4} h\big)_{L^2} \ge 0$$

for all $h \in D\big((-\Delta_R)^{1/2}\big)$ we infer that $(-\Delta_R)^{1/2}$ generates an analytic semigroup of contractions due to Corollary II.4.7 in [23]. Theorem 4.36 in [52] further shows that

$$D\big((-\Delta_R)^{1/2}\big) = \big(L^2(R), H^2(R) \cap H_0^1(R)\big)_{1/2,2}.$$

On the other hand, Δ_R is given by its quadratic form

$$a(u,v) = (\nabla u, \nabla v)_{L^2}$$

on $H_0^1(R)$ and we therefore know due to Theorem VI.2.23 in [47] that $D\big((-\Delta_R)^{1/2}\big) = H_0^1(R)$. So, $\big(L^2(R), H^2(R) \cap H_0^1(R)\big)_{1/2,2}$ is isomorphic to $H_0^1(R)$.

Let $g \in C_c^\infty(\Gamma_1)$ and look at the two restrictions $g_1 \in C_c^\infty(\Gamma_1^-)$ and $g_2 \in C_c^\infty(\Gamma_1^+)$. Let $\chi : [0, a_1^+ - a_1^-] \to \mathbb{R}$ be a C^∞-function with $\operatorname{supp}\chi \subseteq \big[0, \frac{1}{2}(a_1^+ - a_1^-)\big]$ and $\chi = 1$ on $\big[0, \frac{1}{4}(a_1^+ - a_1^-)\big]$. We set

$$\begin{aligned}
&w(x_1, x_2, x_3) \\
&\quad := -\big(\chi(x_1 - a_1^-)(-\Delta_R)^{-1/2} \exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3) \\
&\qquad + \big(\chi(a_1^+ - x_1)(-\Delta_R)^{-1/2} \exp\big((a_1^+ - x_1)(-\Delta_R)^{1/2}\big) g_2\big)(x_2, x_3) \\
&\quad =: w^{(1)}(x_1, x_2, x_3) + w^{(2)}(x_1, x_2, x_3)
\end{aligned}$$

for $(x_1, x_2, x_3) \in \overline{Q}$. By the smoothing of the semigroup, $w(x_1, \cdot, \cdot)$ belongs to $H^2(R)$ for all $x_1 \in Q$. The derivatives of $w^{(1)}$ are given by

$$\begin{aligned}
\partial_1 w^{(1)}(x_1, x_2, x_3) &= -\big(\chi(x_1 - a_1^-) \exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3) \\
&\quad - \big(\chi'(x_1 - a_1^-)(-\Delta_R)^{-1/2} \exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3), \\
\partial_k w^{(1)}(x_1, x_2, x_3) &= -\big(\chi(x_1 - a_1^-)\partial_k(-\Delta_R)^{-1/2} \exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3), \\
\partial_{11} w^{(1)}(x_1, x_2, x_3) &= -\big(\chi(x_1 - a_1^-)(-\Delta_R)^{1/2} \exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3) \\
&\quad - 2\big(\chi'(x_1 - a_1^-) \exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3) \\
&\quad - \big(\chi''(x_1 - a_1^-)(-\Delta_R)^{-1/2} \exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3), \\
\partial_{1k} w^{(1)}(x_1, x_2, x_3) &= -\big(\chi(x_1 - a_1^-)\partial_k \exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3) \\
&\quad - \big(\chi'(x_1 - a_1^-)\partial_k(-\Delta_R)^{-1/2} \exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3), \\
\partial_{kl} w^{(1)}(x_1, x_2, x_3) &= -\big(\chi(x_1 - a_1^-)\partial_{kl}(-\Delta_R)^{-1/2} \exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3)
\end{aligned}$$

for all $k, l \in \{2, 3\}$. Observe that

$$\begin{aligned}
\big\|\exp\big(t(-\Delta_R)^{1/2}\big)h\big\|_{D\left((-\Delta_R)^{1/2}\right)} &= \big\|\exp\big(t(-\Delta_R)^{1/2}\big)h\big\|_{L^2(R)} \\
&\quad + \big\|\exp\big(t(-\Delta_R)^{1/2}\big)(-\Delta_R)^{1/2}h\big\|_{L^2(R)} \\
&\le \|h\|_{D((-\Delta_R)^{1/2})}
\end{aligned}$$

for all $t \ge 0$ and $h \in D\big((-\Delta_R)^{1/2}\big)$. With $g_1 \in D\big((-\Delta_R)^{1/2}\big)$ we infer from Proposition 6.2 in [52] that

$$\begin{aligned}
&\big\|\chi(\cdot - a_1^-)(-\Delta_R)^{1/2} \exp\big((\cdot - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big\|_{L^2}^2 \\
&\quad = \int_{a_1^-}^{a_1^+} \big\|\chi(x_1 - a_1^-)(-\Delta_R)^{1/2} \exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1(x_2, x_3)\big\|_{L^2(R)}^2 \, \mathrm{d}x_1 \\
&\quad \le c \, \|g_1\|^2_{\left(L^2(R), D((-\Delta_R)^{1/2})\right)_{1/2,2}} \cdot (a_1^- - a_1^+) \\
&\quad \le c_* \, \|g_1\|^2_{H_0^{1/2}(R)} .
\end{aligned}$$

We therefore estimate

$$\big\|w^{(1)}\big\|_{L^2}^2 \le \int_{a_1^-}^{a_1^+} \|\chi\|_{L^\infty}^2 \, \big\|(-\Delta_R)^{-1/2}\big\|_{\mathcal{B}(L^2(R))}^2 .$$

$$
\begin{aligned}
&\quad \cdot \left\| \big(\exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3) \right\|_{L^2(R)}^2 \, \mathrm{d}x_1 \\
&\leq c \, \|g_1\|_{L^2(R)}^2 \,, \\
\left\|\partial_1 w^{(1)}\right\|_{L^2}^2 &\leq c \int_{a_1^-}^{a_1^+} \|\chi\|_{L^\infty}^2 \left\| \big(\exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3) \right\|_{L^2(R)}^2 \, \mathrm{d}x_1 \\
&\quad + c \int_{a_1^-}^{a_1^+} \|\chi'\|_{L^\infty}^2 \left\| (-\Delta_R)^{1/2} \right\|_{\mathcal{B}(L^2(R))}^2 \cdot \\
&\qquad \cdot \left\| \big(\exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3) \right\|_{L^2(R)}^2 \, \mathrm{d}x_1 \\
&\leq c \, \|g_1\|_{L^2(R)}^2 \,, \\
\left\|\partial_k w^{(1)}\right\|_{L^2}^2 &\leq c \int_{a_1^-}^{a_1^+} \|\chi\|_{L^\infty}^2 \cdot \\
&\qquad \cdot \left\| \big((-\Delta_R)^{-1/2} \exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3) \right\|_{H_0^1(R)}^2 \, \mathrm{d}x_1 \\
&\leq c \int_{a_1^-}^{a_1^+} \left\| (-\Delta_R)^{-1/2} \right\|_{\mathcal{B}(L^2(R), H_0^1(R))}^2 \|g_1\|_{L^2(R)}^2 \, \mathrm{d}x_1 \\
&\leq c \, \|g_1\|_{L^2(R)}^2 \,, \\
\left\|\partial_{11} w^{(1)}\right\|_{L^2}^2 &\leq c_* \, \|g_1\|_{H_0^{1/2}(R)}^2 + c \, \|g_1\|_{L^2(R)}^2 \,, \\
\left\|\partial_{1k} w^{(1)}\right\|_{L^2}^2 &\leq c \int_{a_1^-}^{a_1^+} \|\chi\|_{L^\infty}^2 \left\| \big((-\Delta_R)^{1/2} \exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3) \right\|_{L^2(R)}^2 \, \mathrm{d}x_1 \\
&\quad + c \int_{a_1^-}^{a_1^+} \|\chi'\|_{L^\infty}^2 \left\| \big(\exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3) \right\|_{L^2(R)}^2 \, \mathrm{d}x_1 \\
&\leq c c_* \, \|g_1\|_{H_0^1(R)}^2 + c \, \|g_1\|_{L^2(R)}^2 \,, \\
\left\|\partial_{kl} w^{(1)}\right\|_{L^2}^2 &\leq c \int_{a_1^-}^{a_1^+} \big\| \chi(x_1 - a_1^-) \big((-\Delta_R)^{-1} (-\Delta_R)^{1/2} \cdot \\
&\qquad \cdot \exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3) \big\|_{H^2(R)}^2 \, \mathrm{d}x_1 \\
&\leq c \int_{a_1^-}^{a_1^+} \left\| (-\Delta_R)^{-1} \right\|_{\mathcal{B}(L^2(R), H^2(R))}^2 \cdot \\
&\qquad \cdot \left\| \chi(x_1 - a_1^-) \big((-\Delta_R)^{1/2} \big(\exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3) \right\|_{H_0^1(R)}^2 \, \mathrm{d}x_1 \\
&\leq c c_* \, \|g_1\|_{H_0^{1/2}(R)} \,,
\end{aligned}
$$

using the equivalence of $\|\cdot\|_{D((-\Delta_R)^{1/2})}$ and $\|\cdot\|_{H^1}$ and the one of $\|\cdot\|_{D(\Delta_R)}$ and $\|\cdot\|_{H^2}$. Together with the analogous estimates for $w^{(2)}$, we derive $w \in H^2(Q)$ and the estimate

$$
\|w\|_{H^2} \leq c \, \|g\|_{H_0^{1/2}(\Gamma_1)} \,.
$$

On Γ_1^- we further obtain

$$
\partial_\nu w(x_1, x_2, x_3)|_{x_1 = a_1^-} = -\partial_1 w^{(1)}(x_1, x_2, x_3)|_{x_1 = a_1^-}
$$

$$= \big(\exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3)|_{x_1 = a_1^-} = g_1(x_2, x_3)$$

and on Γ_1^+

$$\begin{aligned}\partial_\nu w(x_1, x_2, x_3)|_{x_1 = a_1^+} &= \partial_1 w(x_1, x_2, x_3)|_{x_1 = a_1^+} \\ &= \big(\exp\big((a_1^+ - x_1)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3)|_{x_1 = a_1^+} = g_2(x_2, x_3).\end{aligned}$$

The Neumann trace of w on Γ_1 thus equals g. Since $\exp\big(s(-\Delta_R)^{1/2}\big)$ maps into

$$D\big((-\Delta_R)^{1/2}\big) = H_0^1(R)$$

for $s > 0$ and $g_1 \in H_0^1(R)$, the function $w(x_1, \cdot, \cdot)$ has zero trace for all $x_1 \in [a_1^-, a_1^+]$.

We conclude part 2) of the proof with an approximation argument. Let $g \in H_0^{1/2}(\Gamma_1)$ be given. We choose a sequence $(g_n)_{n\in\mathbb{N}}$ in $C_c^\infty(\Gamma_1)$ with $g_n \to g$ in $H_0^{1/2}(\Gamma_1)$ as $n \to \infty$. This is possible since $C_c^\infty(\Gamma_1)$ is dense in $H_0^1(\Gamma_1)$ and thus also in the interpolation space $H_0^{1/2}(\Gamma_1)$, see Proposition 1.17 in [52]. We define the corresponding sequence $(w_n)_{n\in\mathbb{N}}$ in $H^2(Q)$, and with the same estimates as above for w we see

$$\|w_n - w_m\|_{H^2} \le c \, \|g_n - g_m\|_{H_0^{1/2}(\Gamma_1)} \longrightarrow 0$$

as $n, m \to \infty$. Thus, (w_n) has a limit $w \in H^2(Q)$. The continuity of the Dirichlet trace map and the Neumann trace map yields $w = 0$ on $\Gamma \setminus \Gamma_1$ and $\partial_\nu w = g$ on Γ_1.

3) Set $\widetilde{f} := f - w + \Delta w \in L^2(Q)$ with the w from step 2). Part (a) then provides a function $u \in D_0$ with $u - \Delta u = \widetilde{f}$. Hence, $v := u + w \in D$ satisfies $v - \Delta v = f$ and $\partial_\nu v = g$ on $\widetilde{\Gamma}_1'$. By the divergence theorem one checks that v also satisfies (7.8) for all $\varphi \in H^1_{\widetilde{\Gamma}}(Q)$, so that it is equal to $\widetilde{v}$ from step 1). □

We continue with a lemma concerning traces that we need for the trace properties of the Maxwell operators and later on for versions of the splitting operators in an H^1- and an H^2-setting.

Lemma 7.10. *Let $j, k \in \{1, 2, 3\}$ with $k \neq j$.*

(a) For a function $f \in L^2(Q)$ with $\partial_j f, \partial_k f, \partial_{jk} f \in L^2(Q)$ and $f = 0$ on Γ_j we have $\partial_k f = 0$ on Γ_j.

(b) Let $f \in L^2(Q)$ with $\partial_j f \in L^2(Q)$ and $f = 0$ in Γ_j. For $\rho^{(k)} \in C_c^\infty((a_k^-, a_k^+))$ we define the convolution $g := \rho^{(k)} * f$ acting on the k-th variable by extending $\rho^{(k)}$ and f by 0 outside of (a_k^-, a_k^+). Then $g = 0$ on Γ_j.

(c) Let $l \in \mathbb{N}$ be such that $l \ge \frac{1}{a_k^+ - a_k^-}$. Let $f \in L^2(Q)$ with $\partial_j f \in L^2(Q)$ and $f = 0$ on Γ_j. Then

$$g(x) := \int_{a_k^-}^{x_k} \chi_l^{(k)}(t) f(t, \widehat{x}) \, \mathrm{d}t \qquad \text{and} \qquad h(x) := \int_{a_k^-}^{x_k} f(t, \widehat{x}) \, \mathrm{d}t$$

satisfy $g = 0$ *and* $h = 0$ *on* Γ_j*, where* $\chi_l^{(k)}$ *are the cut-off functions defined in* (7.13) *and* $\widehat{x}$ *contains* x_j *and that* x_i *with* $i \in \{1,2,3\} \setminus \{j,k\}$.

PROOF:
Let without loss of generality $j = 1$.
(a) Let $k \in \{2,3\}$ be fixed and recall from (7.11) the set

$$Q_1 := (a_2^-, a_2^+) \times (a_3^-, a_3^+).$$

We have for almost all $(x_2, x_3) \in Q_1$ that $f(\cdot, x_2, x_3) \in H_0^1(a_1^-, a_1^+)$ and

$$f(x_1, x_2, x_3) = \int_{a_1^-}^{x_1} \partial_1 f(t, x_2, x_3) \,\mathrm{d}t,$$

as well as $\partial_{1k} f(x_1, \cdot, \cdot) \in L^2(Q_1)$ for almost all $x_1 \in [a_1^-, a_1^+]$. Let $\varphi \in C_c^\infty(Q)$. Fubini's theorem and integration by parts yields

$$\begin{aligned}
\int_Q f(x)\partial_k\varphi(x)\,\mathrm{d}x &= \int_{a_1^-}^{a_1^+} \int_{Q_1} \int_{a_1^-}^{x_1} \partial_1 f(t,x_2,x_3)\,\mathrm{d}t\ \partial_k\varphi(x_1,x_2,x_3)\,\mathrm{d}(x_2,x_3)\,\mathrm{d}x_1 \\
&= \int_{a_1^-}^{a_1^+} \int_{a_1^-}^{x_1} \int_{Q_1} \partial_1 f(t,x_2,x_3)\partial_k\varphi(x_1,x_2,x_3)\,\mathrm{d}(x_2,x_3)\,\mathrm{d}t\,\mathrm{d}x_1 \\
&= -\int_{a_1^-}^{a_1^+} \int_{a_1^-}^{x_1} \int_{Q_1} \partial_{1k} f(t,x_2,x_3)\varphi(x_1,x_2,x_3)\,\mathrm{d}(x_2,x_3)\,\mathrm{d}t\,\mathrm{d}x_1 \\
&= -\int_Q \int_{a_1^-}^{x_1} \partial_{1k} f(t,x_2,x_3)\,\mathrm{d}t\ \varphi(x)\,\mathrm{d}x.
\end{aligned}$$

This implies $\partial_k f(x_1,x_2,x_3) = \int_{a_1^-}^{x_1} \partial_{1k} f(t,x_2,x_3)\,\mathrm{d}t$ for almost all $(x_2,x_3) \in Q_1$, so that we first get

$$\begin{aligned}
\|\partial_k f(x_1,\cdot,\cdot)\|_{L^2(Q_1)} &\le \int_{a_1^-}^{x_1} \|\partial_{1k} f(t,\cdot,\cdot)\|_{L^2(Q_1)}\,\mathrm{d}t \\
&\le (x_1 - a_1^-)^{1/2} \Big(\int_{a_1^-}^{x_1} \int_{Q_1} |\partial_{1k} f(t,x_2,x_3)|^2 \,\mathrm{d}(x_2,x_3)\,\mathrm{d}t \Big)^{1/2}
\end{aligned} \tag{7.9}$$

for almost all $x_1 \in (a_1^-, a_1^+)$ and then

$$\|\partial_k f(x_1,\cdot,\cdot)\|_{L^2(Q_1)} \le (x_1 - a_1^-)^{1/2} \|\partial_{1k} f\|_{L^2} \longrightarrow 0$$

as $x_1 \to a_1^-$. In the same way as (7.9) we see

$$\|\partial_k f(x_1,\cdot,\cdot)\|_{L^2(Q_1)} \le (a_1^+ - x_1)^{1/2} \Big(\int_{x_1}^{a_1^+} \int_{Q_1} |\partial_{1k} f(t,x_2,x_3)|^2 \,\mathrm{d}(x_2,x_3)\,\mathrm{d}t \Big)^{1/2} \tag{7.10}$$

for almost all $x_1 \in (a_1^-, a_1^+)$.

For $j \in \{1,2,3\}$ we define

$$Q_j := (a_k^-, a_k^+) \times (a_l^-, a_l^+) \tag{7.11}$$

for $k,l \in \{1,2,3\}$ with $k \neq j$, $l \neq j$ and $k \neq l$. For all $n > \frac{4}{d_{\min}}$ we define the set

$$A_n^{(j)} := A_n^{(j),-} \cup A_n^{(j),+} := [a_j^- + \tfrac{1}{n}, a_j^- + \tfrac{2}{n}] \cup [a_j^+ - \tfrac{2}{n}, a_j^+ - \tfrac{1}{n}] \tag{7.12}$$

and the cut-off function

$$\chi_n^{(j)}(t) := \begin{cases} 0, & t \in [a_j^-, a_j^- + \frac{1}{n}], \\ n\big(t - (a_j^- + \frac{1}{n})\big), & t \in (a_j^- + \frac{1}{n}, a_j^- + \frac{2}{n}), \\ 1, & t \in [a_j^- + \frac{2}{n}, a_j^+ - \frac{2}{n}], \\ 1 - n\big(t - (a_j^+ - \frac{2}{n})\big), & t \in (a_j^+ - \frac{2}{n}, a_j^+ - \frac{1}{n}), \\ 0, & t \in [a_j^+ - \frac{1}{n}, a_j^+]. \end{cases} \tag{7.13}$$

Let

$$f_n(t, x_2, x_3) := \chi_n^{(1)}(t) f(t, x_2, x_3).$$

The convergence $\partial_k f_n = \chi_n^{(1)} \partial_k f \to \partial_k f$ in $L^2(Q)$ as $n \to \infty$ is seen with the theorem of dominated convergence. We have the identity

$$\partial_{1k} f_n = (\chi_n^{(1)})' \partial_k f + \chi_n^{(1)} \partial_{1k} f.$$

Using the inequalities (7.9) and (7.10) we deduce

$$\begin{aligned} \big\|(\chi_n^{(1)})' \partial_k f\big\|_{L^2} &\leq \Big(\int_{A_n^{(1)}} \int_{Q_1} n^2 \left|\partial_k f(x_1, x_2, x_3)\right|^2 \,\mathrm{d}(x_2, x_3)\,\mathrm{d}x_1 \Big)^{1/2} \\ &\leq \Big(2n \sup_{x_1 \in A_n^{(1)}} \int_{Q_1} \left|\partial_k f(x_1, x_2, x_3)\right|^2 \,\mathrm{d}(x_2, x_3) \Big)^{1/2} \\ &\leq \Big(4 \sup_{x_1 \in A_n^{(1),-}} \int_{a_1^-}^{x_1} \int_{Q_1} \left|\partial_{1k} f(t, x_2, x_3)\right|^2 \,\mathrm{d}(x_2, x_3)\,\mathrm{d}t \\ &\qquad + 4 \sup_{x_1 \in A_n^{(1),+}} \int_{x_1}^{a_1^+} \int_{Q_1} \left|\partial_{1k} f(t, x_2, x_3)\right|^2 \,\mathrm{d}(x_2, x_3)\,\mathrm{d}t \Big)^{1/2} \\ &= 2 \Big(\int_{[a_1^-, a_1^- + \frac{2}{n}] \cup [a_1^+ - \frac{2}{n}, a_1^+]} \int_{Q_1} \left|\partial_{1k} f(t, x_2, x_3)\right|^2 \,\mathrm{d}(x_2, x_3)\,\mathrm{d}t \Big)^{1/2} \longrightarrow 0 \end{aligned}$$

as $n \to \infty$, where we used the theorem of dominated convergence in the last step. Together with $\chi_n^{(1)} \partial_{1k} f \to \partial_{1k} f$ in $L^2(Q)$ as $n \to \infty$ it follows $\partial_{1k} f_n \to \partial_{1k} f$ in $L^2(Q)$ as $n \to \infty$. We conclude that on Γ_1 the trace of $\partial_k f_n$ converges to the trace of $\partial_k f$ in the L^2-sense. Since $\partial_k f_n = 0$ on Γ_1 is evident due to the cut-off function, the claim follows.

(b) We define

$$f_n(x_1, x_2, x_3) := \chi_n^{(1)}(x_1) f(x_1, x_2, x_3)$$

From $\chi_n^{(1)} f \to f$ in $L^2(Q)$ we infer $\rho^{(k)} * f_n \to \rho * f$ in $L^2(Q)$ as $n \to \infty$. The convergence

$$\partial_1\big(\rho^{(k)} * f_n\big) = \rho^{(k)} * \big((\chi_n^{(1)})' f + \chi_n^{(1)} \partial_1 f\big) \longrightarrow \rho * \partial_1 f$$

as $n \to \infty$ is seen with the methods of the proof of part (a). As above, we thus get the claim due to $f_n = 0$ in a neighbourhood of Γ_1 and thus $\rho^{(k)} * f_n = 0$ on Γ_1.

(c) Let without loss of generality $k = 2$. We define

$$g_n(x_1, x_2, x_3) := \chi_n^{(1)}(x_1) \int_{a_2^-}^{x_2} \chi_l^{(2)}(t) f(x_1, t, x_3)\,\mathrm{d}t.$$

From

$$\left\| \chi_n^{(1)}(x_1) \int_{a_2^-}^{x_2} \chi_l^{(2)}(t) f(x_1, s, x_3)\,\mathrm{d}t \right\|_{L^2} \leq (a_2^+ - a_2^-)\, \|f\|_{L^2}$$

we infer $g_n \to g$ in $L^2(Q)$ as $n \to \infty$ with the theorem of dominated convergence. We compute

$$\begin{aligned} \partial_1 g_n(x_1, x_2, x_3) = {} & (\chi_n^{(1)})'(x_1) \int_{a_2^-}^{x_2} \chi_l^{(2)}(t) f(x_1, t, x_3)\,\mathrm{d}t \\ & + \chi_n^{(1)}(x_1) \int_{a_2^-}^{x_2} \chi_l^{(2)}(t) \partial_1 f(x_1, t, x_3)\,\mathrm{d}t \end{aligned}$$

and see

$$\partial_1 g_n \longrightarrow \int_{a_2^-}^{x_2} \chi_l^{(2)}(t) \partial_1 f(x_1, t, x_3)\,\mathrm{d}t$$

as $n \to \infty$ with the methods of the proof of part (a). As above, we thus get the claim due to $g_n = 0$ on Γ_1. The proof for h is done in the same way. □

We now are in the position to prove embedding and trace properties of the domains of the Maxwell operators. In Lemma 3.2 in [37] it was shown that $D(M_0^2)$ is continuously embedded into $H^2(Q)$. More results on embeddings and traces on Lipschitz domains can be found in great detail in [15].

Proposition 7.11. *(a) The domain $D(M_{\mathrm{div}}^{(0)})$ is continuously embedded into $H^1(Q)^6$. Furthermore, we have*

$$\|(\boldsymbol{E}, \boldsymbol{H})\|_{H^1} \leq c\big(\|(\boldsymbol{E}, \boldsymbol{H})\|_{X_{\mathrm{div}}^{(0)}} + \|M(\boldsymbol{E}, \boldsymbol{H})\|_X\big)$$

for all $(\boldsymbol{E}, \boldsymbol{H}) \in D(M_{\mathrm{div}}^{(0)})$ with a constant depending only on $\|\varepsilon\|_{W^{1,\infty}}$, $\|\mu\|_{W^{1,\infty}}$, $\|\sigma\|_{L^\infty}$ and δ. Additionally, $(\boldsymbol{E}, \boldsymbol{H}) \in D(M_{\mathrm{div}}^{(0)})$ has the traces

$$\begin{aligned} E_2 = E_3 = 0, \qquad & H_1 = 0 \qquad \text{on } \Gamma_1, \\ E_1 = E_3 = 0, \qquad & H_2 = 0 \qquad \text{on } \Gamma_2, \\ E_1 = E_2 = 0, \qquad & H_3 = 0 \qquad \text{on } \Gamma_3. \end{aligned}$$

(b) Let $\varepsilon, \mu \in W^{2,3}(Q)$. Then $X^{(2)}_{\mathrm{div}}$ is continuously embedded into $H^2(Q)^6$. Moreover, we have

$$\|(\boldsymbol{E}, \boldsymbol{H})\|_{H^2} \le c\,\|(\boldsymbol{E}, \boldsymbol{H})\|_{X^{(2)}_{\mathrm{div}}}$$

for all $(\boldsymbol{E}, \boldsymbol{H}) \in X^{(2)}_{\mathrm{div}}$, where the constants depends only on the quantities $\|\varepsilon\|_{W^{1,\infty}\cap W^{2,3}}$, $\|\mu\|_{W^{1,\infty}\cap W^{2,3}}$, $\|\sigma\|_{W^{1,\infty}}$ and δ. Additionally, $(\boldsymbol{E}, \boldsymbol{H}) \in X^{(2)}_{\mathrm{div}}$ has the traces

$$\begin{aligned}
E_2 = E_3 = 0, &\qquad \partial_2 E_2 = \partial_3 E_2 = \partial_2 E_3 = \partial_3 E_3 = 0 &&\text{on } \Gamma_1,\\
E_1 = E_3 = 0, &\qquad \partial_1 E_1 = \partial_3 E_1 = \partial_1 E_3 = \partial_3 E_3 = 0 &&\text{on } \Gamma_2,\\
E_1 = E_2 = 0, &\qquad \partial_1 E_1 = \partial_2 E_1 = \partial_1 E_2 = \partial_2 E_2 = 0 &&\text{on } \Gamma_3,\\
H_1 = 0, &\qquad \partial_2 H_1 = \partial_3 H_1 = 0 &&\text{on } \Gamma_1,\\
H_2 = 0, &\qquad \partial_1 H_2 = \partial_3 H_2 = 0 &&\text{on } \Gamma_2,\\
H_3 = 0, &\qquad \partial_1 H_3 = \partial_2 H_3 = 0 &&\text{on } \Gamma_3.
\end{aligned}$$

Proof:
(a) Let $(\mathbf{E}, \mathbf{H}) \in D(M^{(0)}_{\mathrm{div}})$. The embedding is a consequence of Lemma 7.7 and 7.5 and furthermore Proposition 7.4 since $\operatorname{curl}\mathbf{H} = \sigma\mathbf{E} + \varepsilon\big(M(\mathbf{E}, \mathbf{H})\big)_1$. The trace result is clear if $(\mathbf{E}, \mathbf{H})$ is also smooth in $\overline{Q}$, and this follows by an approximation, using Proposition 7.3.

(b) Let $(\mathbf{E}, \mathbf{H}) \in X^{(2)}_{\mathrm{div}}$.

1) By part (a) and $X^{(2)}_{\mathrm{div}} \hookrightarrow D(M^{(0)}_{\mathrm{div}})$, $\mathbf{E}$ and $\mathbf{H}$ belong to $H^1(Q)^3$ and satisfy the assertion on the zero-order traces. We next show that both fields are contained in $H^2_{loc}(Q)^3$. Part (a) provides the estimate

$$\|(\mathbf{E}, \mathbf{H})\|_{H^1} \le c\big(\|(\mathbf{E}, \mathbf{H})\|_{X^{(0)}_{\mathrm{div}}} + \|M(\mathbf{E}, \mathbf{H})\|_X\big). \tag{7.14}$$

The momentum inequality, see Theorem II.5.34 in [23], together with Young's inequality yields

$$\|M(\mathbf{E}, \mathbf{H})\|_X \le c\big(\|(\mathbf{E}, \mathbf{H})\|_X + \big\|M^2(\mathbf{E}, \mathbf{H})\big\|_X\big). \tag{7.15}$$

We have

$$M^2(\mathbf{E}, \mathbf{H}) = \begin{pmatrix} \frac{\sigma^2}{\varepsilon^2}\mathbf{E} - \frac{1}{\varepsilon}\operatorname{curl}\frac{1}{\mu}\operatorname{curl}\mathbf{E} - \frac{\sigma}{\varepsilon^2}\operatorname{curl}\mathbf{H} \\ \frac{1}{\mu}\nabla\big(\frac{\sigma}{\varepsilon}\big) \times \mathbf{E} + \frac{\sigma}{\mu\varepsilon}\operatorname{curl}\mathbf{E} - \frac{1}{\mu}\operatorname{curl}\frac{1}{\varepsilon}\operatorname{curl}\mathbf{H} \end{pmatrix} =: \begin{pmatrix} \big(M^2(\mathbf{E}, \mathbf{H})\big)_1 \\ \big(M^2(\mathbf{E}, \mathbf{H})\big)_2 \end{pmatrix}$$

and furthermore, due to identity (7.5),

$$\begin{aligned}
\operatorname{curl}\big(\tfrac{1}{\mu}\operatorname{curl}\mathbf{E}\big) &= \big(\nabla\tfrac{1}{\mu}\big) \times \operatorname{curl}\mathbf{E} + \frac{1}{\mu}\operatorname{curl}\operatorname{curl}\mathbf{E}\\
&= -\frac{1}{\mu^2}(\nabla\mu) \times \operatorname{curl}\mathbf{E} + \frac{1}{\mu}(-\Delta\mathbf{E} + \nabla\operatorname{div}\mathbf{E})\\
&= -\frac{1}{\mu^2}(\nabla\mu) \times \operatorname{curl}\mathbf{E} - \frac{1}{\mu}\Delta\mathbf{E} + \frac{1}{\mu}\nabla\big(\tfrac{1}{\varepsilon}\operatorname{div}(\varepsilon\mathbf{E}) - \tfrac{1}{\varepsilon}\nabla\varepsilon \cdot \mathbf{E}\big)
\end{aligned}$$

$$= -\frac{1}{\mu^2}(\nabla\mu)\times\operatorname{curl}\mathbf{E} - \frac{1}{\mu}\Delta\mathbf{E} - \frac{1}{\mu\varepsilon^2}\operatorname{div}(\varepsilon\mathbf{E})\nabla\varepsilon + \frac{1}{\mu\varepsilon}\nabla\operatorname{div}(\varepsilon\mathbf{E}) + \frac{1}{\mu\varepsilon^2}(\nabla\varepsilon\cdot\mathbf{E})\nabla\varepsilon - \frac{1}{\mu\varepsilon}\nabla(\nabla\varepsilon\cdot\mathbf{E}).$$

Reordering these terms gives

$$\begin{aligned}\Delta\mathbf{E} = \mu\varepsilon\big(M^2(\mathbf{E},\mathbf{H})\big)_1 - \frac{\mu\sigma^2}{\varepsilon}\mathbf{E} + \frac{\mu\sigma}{\varepsilon}\operatorname{curl}\mathbf{H} \\ - \frac{1}{\mu}(\nabla\mu)\times\operatorname{curl}\mathbf{E} - \frac{1}{\varepsilon^2}\operatorname{div}(\varepsilon\mathbf{E})\nabla\varepsilon + \frac{1}{\varepsilon}\nabla\operatorname{div}(\varepsilon\mathbf{E}) \\ + \frac{1}{\varepsilon^2}(\nabla\varepsilon\cdot\mathbf{E})\nabla\varepsilon - \frac{1}{\varepsilon}\bigg(\sum_{j=1}^{3}\big((\partial_{jk}\varepsilon)E_j + (\partial_j\varepsilon)\partial_k E_j\big)\bigg)_{k=1}^{3}.\end{aligned}$$

We now use $(M^2(\mathbf{E},\mathbf{H}))_1 \in L^2(Q)^3$, (7.14), $\operatorname{div}(\varepsilon\mathbf{E}) \in H^1(Q)$, the Sobolev embedding $H^1(Q)^3 \hookrightarrow L^6(Q)^3$ applied to $\mathbf{E}$, the assumptions on ε, μ and σ and (7.15). It follows

$$\|\Delta\mathbf{E}\|_{L^2} \le c\big(\big\|M^2(\mathbf{E},\mathbf{H})\big\|_X + \|(\mathbf{E},\mathbf{H})\|_{X^{(0)}_{\operatorname{div}}}\big).$$

Here, c depends only on $\|\varepsilon\|_{W^{1,\infty}\cap W^{2,3}}$, $\|\mu\|_{W^{1,\infty}}$, $\|\sigma\|_{L^\infty}$ and δ. So, ΔE_j belongs to $L^2(Q)$ for all $j \in \{1,2,3\}$. Analogously, we infer with $\operatorname{div}(\mu\mathbf{H}) = 0$ first

$$\begin{aligned}\operatorname{curl}\big(\tfrac{1}{\varepsilon}\operatorname{curl}\mathbf{H}\big) = -\frac{1}{\varepsilon^2}(\nabla\varepsilon)\times\operatorname{curl}\mathbf{H} - \frac{1}{\varepsilon}\Delta\mathbf{H} \\ + \frac{1}{\varepsilon^2\mu}(\nabla\mu\cdot\mathbf{H})\nabla\mu - \frac{1}{\varepsilon\mu}\nabla(\nabla\mu\cdot\mathbf{H})\end{aligned}$$

and then

$$\begin{aligned}\Delta\mathbf{H} = \varepsilon\mu\big(M^2(\mathbf{E},\mathbf{H})\big)_2 - \varepsilon\nabla\big(\tfrac{\sigma}{\varepsilon}\big)\times\mathbf{E} - \sigma\operatorname{curl}\mathbf{E} \\ - \frac{1}{\varepsilon}(\nabla\varepsilon)\times\operatorname{curl}\mathbf{H} + \frac{1}{\varepsilon\mu}(\nabla\mu\cdot\mathbf{H})\nabla\mu - \frac{1}{\mu}\nabla(\nabla\mu\cdot\mathbf{E}).\end{aligned}$$

In the same way as above, this identity yields

$$\|\Delta\mathbf{H}\|_{L^2} \le c\big(\big\|M^2(\mathbf{E},\mathbf{H})\big\|_X + \|(\mathbf{E},\mathbf{H})\|_{X^{(0)}_{\operatorname{div}}}\big)$$

and $\Delta H_j \in L^2(Q)$ for all $j \in \{1,2,3\}$, where now c also depends on $\|\mu\|_{W^{2,3}}$ and $\|\sigma\|_{W^{1,\infty}}$. So, we have shown $(\Delta\mathbf{E},\Delta\mathbf{H}) \in L^2(Q)^6$ and

$$\|(\Delta\mathbf{E},\Delta\mathbf{H})\|_{L^2} \le c\big(\big\|M^2(\mathbf{E},\mathbf{H})\big\|_X + \|(\mathbf{E},\mathbf{H})\|_{X^{(0)}_{\operatorname{div}}}\big). \tag{7.16}$$

We recall the definition

$$H^2_{loc}(Q) := \{u \in L^1_{loc}(Q) \mid u \in H^2(U) \text{ for each open set } U \subseteq \overline{U} \subseteq Q\}.$$

Let $Q_0 \subseteq Q$ be open and U a domain with $Q_0 \subseteq U \subseteq \overline{U} \subseteq Q$. Let $\varphi \in C_c^\infty(Q)$ with $\varphi = 1$ on U and $u \in \{E_j, H_j, j \in \{1,2,3\}\}$. The function $v := \varphi u$ satisfies $v \in H_0^1(Q)$ and $v = u$ on U. Since $\mathbf{E}, \mathbf{H} \in H^1(Q)^3$, the function v belongs to $H^1(U)$. From the identity

$$\Delta v = u\Delta\varphi + 2\nabla u \cdot \nabla\varphi + \varphi\Delta u$$

and $\Delta E_j, \Delta H_j \in L^2(Q)$ for all $j \in \{1,2,3\}$ we deduce that Δv belongs to $L^2(U)$. Theorem 8.8 in [31] then implies that $v \in H^2(U)$ and hence $u \in H^2(Q_0)$. As a result, E_j and H_j are contained in $H^2_{loc}(Q)$ for all $j \in \{1,2,3\}$.

2) We next show $\mathbf{E} \in H^2(Q)^3$ and the estimate $\|\mathbf{E}\|_{H^2} \leq c\,\|(\mathbf{E},\mathbf{H})\|_{X^{(2)}_{\mathrm{div}}}$. For this part of the proof we set $\widetilde{\Gamma} := \Gamma_2 \cup \Gamma_3$. Observe that

$$\Delta(\varepsilon E_1) = E_1\Delta\varepsilon + 2\nabla\varepsilon \cdot \nabla E_1 + \varepsilon\Delta E_1.$$

From $\Delta E_1 \in L^2(Q)$, $E_1 \in H^1(Q)$, the assumption on ε and the embedding $H^1(Q) \hookrightarrow L^6(Q)$, we thus conclude that $(I - \Delta)\varepsilon E_1$ belongs to $L^2(Q)$. We further compute that

$$\partial_{kl}(\varepsilon E_1) = E_1\partial_{lk}\varepsilon + (\partial_k\varepsilon)(\partial_l E_1) + (\partial_l\varepsilon)(\partial_k E_1) + \varepsilon\partial_{kl}E_1 \tag{7.17}$$

for all $k, l \in \{1,2,3\}$. Using $E_1 \in H^2_{loc}(Q)$ and $E_1 \in H^1(Q)$, we infer that εE_1 is contained in $H^2_{loc}(Q)$. Lemma 7.5 and $E_1 = 0$ on $\Gamma_2 \cup \Gamma_3$ show that $\varepsilon E_1 = 0$ on $\Gamma_2 \cup \Gamma_3$. We fix a function $\psi \in H^1(Q)$ with $\partial_2\psi, \partial_3\psi \in H^1(Q)$ and essential support in

$$Q^{(\eta)} := [a_1^-, a_1^+] \times [a_2^- + \eta, a_2^+ - \eta] \times [a_3^- + \eta, a_3^+ - \eta] \tag{7.18}$$

for an $\eta = \eta(\psi) \in (0, d_{\min}/2)$. For each $\kappa \in (0, d_{\min}/2)$ we define

$$Q_\kappa := (a_1^- + \kappa, a_1^+ - \kappa) \times (a_2^- + \kappa, a_2^+ - \kappa) \times (a_3^- + \kappa, a_3^+ - \kappa).$$

We take $\kappa \in (0, \eta)$ and denote by $\Gamma_1^\pm(\kappa)$ those open faces of Q_κ that contain the points of the form $(a_1^\mp \pm \kappa, x_2, x_3)$. We conclude with the theorem of dominated convergence, integration by parts and $\nabla(\varepsilon E_j) \in H^1(Q)$ that

$$\begin{aligned}
\int_Q \varepsilon E_1\psi\,\mathrm{d}x + \int_Q \nabla(\varepsilon E_1)\cdot\nabla\psi\,\mathrm{d}x &= \lim_{\kappa\to 0}\int_{Q_\kappa}\big(\varepsilon E_1\psi + \nabla(\varepsilon E_1)\cdot\nabla\psi\big)\,\mathrm{d}x \\
&= \lim_{\kappa\to 0}\left[\int_{Q_\kappa}\psi(I-\Delta)(\varepsilon E_1)\,\mathrm{d}x + \int_{\partial Q_\kappa}\mathrm{tr}_n\big(\psi\nabla(\varepsilon E_1)\big)\,\mathrm{d}\sigma\right] \\
&= \int_Q \psi(I-\Delta)(\varepsilon E_1)\,\mathrm{d}x \pm \lim_{\kappa\to 0}\int_{\Gamma_1^\pm(\kappa)}\mathrm{tr}_n\big(\psi\nabla(\varepsilon E_1)\big)\,\mathrm{d}\sigma.
\end{aligned}$$

Moreover, the boundary of $\Gamma_1^\pm(\kappa)$ is disjoint to $Q^{(\eta)}$ due to $\kappa < \eta$. Hence, ψ vanishes on the boundary of $\Gamma_1^\pm(\kappa)$, so that

$$\pm\int_{\Gamma_1^\pm(\kappa)}\psi\big(\partial_2(\varepsilon E_2) + \partial_3(\varepsilon E_3)\big)\,\mathrm{d}\sigma = \mp\int_{\Gamma_1^\pm(\kappa)}\big(\varepsilon E_2\partial_2\psi + \varepsilon E_3\partial_3\psi\big)\,\mathrm{d}\sigma.$$

Therefore, we can continue our calculation by

$$\begin{aligned}
&\int_Q \varepsilon E_1 \psi \,\mathrm{d}x + \int_Q \nabla(\varepsilon E_1)\cdot\nabla\psi \,\mathrm{d}x \\
&\quad= \int_Q \psi(I-\Delta)(\varepsilon E_1)\,\mathrm{d}x \pm \lim_{\kappa\to 0}\int_{\Gamma_1^{\pm}(\kappa)} \psi\partial_1(\varepsilon E_1)\,\mathrm{d}\sigma \\
&\quad= \int_Q \psi(I-\Delta)(\varepsilon E_1)\,\mathrm{d}x \pm \lim_{\kappa\to 0}\int_{\Gamma_1^{\pm}(\kappa)} \psi\,\mathrm{div}(\varepsilon E_1)\,\mathrm{d}\sigma \\
&\qquad\mp \lim_{\kappa\to 0}\int_{\Gamma_1^{\pm}(\kappa)} \psi\big(\partial_2(\varepsilon E_2)+\partial_3(\varepsilon E_3)\big)\,\mathrm{d}\sigma \\
&\quad= \int_Q \psi(I-\Delta)(\varepsilon E_1)\,\mathrm{d}x + \lim_{\kappa\to 0}\int_{\partial Q_\kappa} \psi\rho\,\mathrm{d}\sigma \\
&\qquad\pm \lim_{\kappa\to 0}\int_{\Gamma_1^{\pm}(\kappa)} \big(\varepsilon E_2\partial_2\psi + \varepsilon E_3\partial_3\psi\big)\,\mathrm{d}\sigma.
\end{aligned}$$

Lemma 7.5 together with part 1) implies $\varepsilon E_j = 0$ on Γ_1 for all $j \in \{2,3\}$, so that Lemma 7.8 yields

$$\lim_{\kappa\to 0}\int_{\partial Q_\kappa} \psi\rho\,\mathrm{d}\sigma = \int_\Gamma \psi\rho\,\mathrm{d}\sigma \qquad \text{and} \qquad \lim_{\kappa\to 0}\int_{\Gamma_1^{\pm}(\kappa)} \big(\varepsilon E_2\partial_2\psi + \varepsilon E_3\partial_3\psi\big)\,\mathrm{d}\sigma = 0.$$

We thus have shown

$$\begin{aligned}
&\int_Q \varepsilon E_1\psi\,\mathrm{d}x + \int_Q \nabla(\varepsilon E_1)\cdot\nabla\psi\,\mathrm{d}x \\
&\qquad= \int_Q \psi(I-\Delta)(\varepsilon E_1)\,\mathrm{d}x + \int_\Gamma \psi\rho\,\mathrm{d}\sigma.
\end{aligned} \tag{7.19}$$

We next show that we can approximate each function in $H^1_{\Gamma_2\cup\Gamma_3}(Q)$ in $H^1(Q)$ by functions as chosen above. Let $\psi \in H^1_{\Gamma_2\cup\Gamma_3}(Q)$ and $\eta > 0$. Take functions $\widetilde{\varphi}_m \in C^\infty(\overline{Q})$ with $\widetilde{\varphi}_m \to \psi$ in $H^1(Q)$ as $m\to\infty$. Then $\mathrm{tr}(\widetilde{\varphi}_m) \to \mathrm{tr}(\psi) = 0$ in $L^2(\Gamma_2\cup\Gamma_3)$. We fix an $m\in\mathbb{N}$ with

$$\|\widetilde{\varphi}_m - \psi\|_{H^1} \le \eta \qquad \text{and} \qquad \|\mathrm{tr}(\widetilde{\varphi}_m)\|_{L^2(\Gamma_2\cup\Gamma_3)} \le \eta.$$

Set $\widetilde{\varphi} := \widetilde{\varphi}_m$. Recall for all $n > \frac{4}{d_{\min}}$ the sets $A_n^{(j)}$, $A_n^{(j),+}$ and $A_n^{(j),-}$ in (7.12) and the cut-off function in (7.13). Set $\varphi_n := \chi_n^{(2)}\chi_n^{(3)}\widetilde{\varphi}$. The theorem of dominated convergence gives $\varphi_n \to \widetilde{\varphi}$ and $\partial_1\varphi_n \to \partial_1\widetilde{\varphi}$ in $L^2(Q)$ as $n\to\infty$. Additionally, we get with

$$\begin{aligned}
\widetilde{\varphi}(x_1,x_2,x_3) &= \int_{a_2^-}^{x_2} \partial_1\widetilde{\varphi}(x_1,t,x_3)\,\mathrm{d}t - \widetilde{\varphi}(x_1,a_2^-,x_3), \\
\widetilde{\varphi}(x_1,x_2,x_3) &= -\int_{x_2}^{a_2^+} \partial_1\widetilde{\varphi}(x_1,t,x_3)\,\mathrm{d}t + \widetilde{\varphi}(x_1,a_2^+,x_3)
\end{aligned}$$

that

$$\Big\|\big(\chi_n^{(2)}\big)'\chi_n^{(3)}\widetilde{\varphi}\Big\|_{L^2}^2 \le \int_{A_n^{(2)}}\int_{Q_2} n^2\,|\widetilde{\varphi}(x_1,x_2,x_3)|^2\,\mathrm{d}(x_1,x_3)\,\mathrm{d}x_2$$

$$
\begin{aligned}
&\leq 2n \sup_{x_2 \in A_n^{(2)}} \int_{Q_2} |\widetilde{\varphi}(x_1,x_2,x_3)|^2 \, \mathrm{d}(x_1,x_3) \\
&\leq cn \sup_{x_2 \in A_n^{(2),-}} \int_{a_2^-}^{x_2} \int_{Q_2} |\partial_1 \widetilde{\varphi}(x_1,t,x_3)|^2 \, \mathrm{d}(x_1,x_3)\,\mathrm{d}t \\
&\quad + cn \sup_{x_2 \in A_n^{(2),-}} \int_{Q_2} \left|\widetilde{\varphi}(x_1,a_2^-,x_3)\right|^2 \mathrm{d}(x_1,x_3)\,\mathrm{d}t \\
&\quad + cn \sup_{x_2 \in A_n^{(2),+}} \int_{x_2}^{a_2^+} \int_{Q_2} |\partial_1 \widetilde{\varphi}(x_1,t,x_3)|^2 \, \mathrm{d}(x_1,x_3)\,\mathrm{d}t \\
&\quad + cn \sup_{x_2 \in A_n^{(2),+}} \int_{Q_2} \left|\widetilde{\varphi}(x_1,a_2^+,x_3)\right|^2 \mathrm{d}(x_1,x_3)\,\mathrm{d}t \\
&\leq c \int_{[a_2^-,a_2^-+\frac{2}{n}]\cup[a_2^+-\frac{2}{n},a_2^+]} \int_{Q_2} |\partial_1 \widetilde{\varphi}(x_1,t,x_3)|^2 \, \mathrm{d}(x_2,x_3)\,\mathrm{d}t + \widetilde{c}\,\|\mathrm{tr}(\widetilde{\varphi})\|_{L^2(\Gamma_2)}^2 \\
&\leq (\widetilde{c}+1)\eta^2
\end{aligned}
$$

for n large enough, since the first summand tends to zero as $n \to \infty$ and $\|\mathrm{tr}(\widetilde{\varphi})\|_{L^2(\Gamma_2)}^2 \leq \eta^2$. Hence,

$$
\partial_2 \varphi_n - \partial_2 \widetilde{\varphi} = \big(\chi_n^{(2)} \chi_n^{(3)} - \mathbb{1}\big)\partial_2 \widetilde{\varphi} + \big(\chi_n^{(2)}\big)' \chi_n^{(3)} \widetilde{\varphi} \longrightarrow 0
$$

in $L^2(Q)$ as $n \to \infty$ by the theorem of dominated convergence. Analogously we see $\partial_2 \varphi_n \to \partial_3 \widetilde{\varphi}$ in $L^2(Q)$ as $n \to \infty$. Altogether, we $\varphi_n \to \widetilde{\varphi}$ in $H^1(Q)$ as $n \to \infty$.

Therefore, (7.19) holds true for all $\psi \in H^1_\Gamma(Q)$ by approximation. Lemma 7.9 shows that εE_1 is contained in $H^2(Q)$ and that

$$
\begin{aligned}
\|\varepsilon E_1\|_{H^2} &\leq c\big(\|\varepsilon E_1 - \Delta(\varepsilon E_1)\|_{L^2} + \|\rho\|_{H_0^{1/2}(\Gamma_1)}\big) \\
&\leq c\big(\big\|M^2(\mathbf{E},\mathbf{H})\big\|_X + \|(\mathbf{E},\mathbf{H})\|_{X_{\mathrm{div}}^{(0)}} + \|\rho\|_{H_0^{1/2}(\Gamma_1)}\big) \\
&\leq c\,\|(\mathbf{E},\mathbf{H})\|_{X_{\mathrm{div}}^{(2)}}\,,
\end{aligned} \tag{7.20}
$$

using estimate (7.16) in the second to the last estimate.

For all $k,l \in \{1,2,3\}$ we have

$$
\partial_{kl} E_1 = \frac{1}{\varepsilon}\partial_{kl}(\varepsilon E_1) - \frac{\partial_k \varepsilon}{\varepsilon^2}\partial_l(\varepsilon E_1) - \frac{\partial_l \varepsilon}{\varepsilon^2}\partial_k(\varepsilon E_1) + \varepsilon E_1\big(-\tfrac{\partial_{lk}\varepsilon}{\varepsilon^2} + \tfrac{2(\partial_k \varepsilon)(\partial_l \varepsilon)}{\varepsilon^3}\big).
$$

Using $E_1 \in H^1(Q) \hookrightarrow L^6(Q)$, $\varepsilon E_1 \in H^2(Q)$ and the assumptions on ε, we conclude from this that E_1 belongs to $H^2(Q)$. E_2 and E_3 are treated in the same way, giving $\mathbf{E} \in H^2(Q)^3$. From (7.20) we infer

$$
\begin{aligned}
\|E_1\|_{H^2}^2 \leq c\Bigg(&\tfrac{1}{\delta^2}\,\|\varepsilon E_1\|_{L^2}^2 + 2\sum_{k=1}^{3}\big(\tfrac{1}{\delta^2}\,\|\partial_k(\varepsilon E_1)\|_{L^2}^2 + \tfrac{1}{\delta^4}\,\|\partial_k \varepsilon\|_{L^\infty}^2\,\|\varepsilon E_1\|_{L^2}^2\big) \\
&+ \sum_{k,l=1}^{3}\Big(\tfrac{1}{\delta^2}\,\|\partial_{kl}(\varepsilon E_1)\|_{L^2}^2 + \tfrac{\|\partial_k \varepsilon\|_{L^\infty}^2}{\delta^4}\,\|\partial_l(\varepsilon E_1)\|_{L^2}^2 + \tfrac{\|\partial_l \varepsilon\|_{L^\infty}^2}{\delta^4}\,\|\partial_k(\varepsilon E_1)\|_{L^2}^2
\end{aligned}
$$

$$+ \Big(\tfrac{\|\partial_{kl}\varepsilon\|_{L^\infty}^2}{\delta^4} + \tfrac{2\|\partial_k\varepsilon\|_{L^\infty}^2 \|\partial_l\varepsilon\|_{L^\infty}^2}{\delta^6} \Big) \|\varepsilon E_1\|_{L^2}^2 \Big) \Big)$$
$$\leq c \|\varepsilon E_1\|_{H^2}^2 \leq c \|(\mathbf{E},\mathbf{H})\|_{X_{\mathrm{div}}^{(2)}}^2 ,$$

which is the desired norm estimate.

3) For $i \in \{1,2,3\}$ we denote by γ_i the Dirichlet trace operator on Γ_i. Let $i,j,k \in \{1,2,3\}$ with $i \neq j$ and $i \neq k$. We approximate the function $E_k \in H^2(Q)$ in $H^2(Q)$ by a sequence $(v_n)_{n\in\mathbb{N}} \subseteq C^2(\overline{Q})$. Observe that $\gamma_i \partial_j v_n = \partial_j \gamma_i v_n$. Taking the limit $n \to \infty$ gives with the continuity of the trace operators that $\gamma_i \partial_j E_k = \partial_j \gamma_i E_k$, so that the already established zero-order traces of $\mathbf{E}$ imply now the claimed first-order traces of $\mathbf{E}$ by Lemma 7.10.

4) Using Lemma 7.9, the remaining assertions for $\mathbf{H}$ can be seen as in the proof of Lemma 3.7 in [37]. □

One benefit of the embeddings we have just seen is that the Maxwell operators map into the respective restrictions of X.

Lemma 7.12. *(a) If $\sigma = 0$, then the operator M_0 maps into X_0 and is thus equal to the part of M in X_0.*

(b) The operator $M_{\mathrm{div}}^{(0)}$ maps into $X_{\mathrm{div}}^{(0)}$ and is thus the part of M in $X_{\mathrm{div}}^{(0)}$.

(c) If $\varepsilon, \mu, \sigma \in W^{2,3}(Q)$, then the operator $M_{\mathrm{div}}^{(2)}$ maps into $X_{\mathrm{div}}^{(2)}$.

Proof:
The proof for part (a) can be found in the proof of Proposition 3.5 in [37].

(b) Let $(\mathbf{E},\mathbf{H}) \in D(M_{\mathrm{div}}^{(0)})$. With $\operatorname{div}\operatorname{curl} = 0$ we compute

$$\Xi := \operatorname{div}\big(\varepsilon(M(\mathbf{E},\mathbf{H}))_1\big) = \operatorname{div}\big(\varepsilon\tfrac{\sigma}{\varepsilon}\mathbf{E}\big) = \nabla\big(\tfrac{\sigma}{\varepsilon}\big)\cdot\varepsilon\mathbf{E} + \frac{\sigma}{\varepsilon}\operatorname{div}(\varepsilon\mathbf{E})$$
$$= \nabla\sigma\cdot\mathbf{E} - \frac{\sigma}{\varepsilon}\nabla\varepsilon\cdot\mathbf{E} + \frac{\sigma}{\varepsilon}\operatorname{div}(\varepsilon\mathbf{E}). \tag{7.21}$$

This function belongs to $L^2(Q)$ due to the general assumptions on ε and σ and since $(\mathbf{E},\mathbf{H}) \in X_{\mathrm{div}}^{(0)}$. The statement for $X_{\mathrm{div}}^{(0)}$ now follows as the one for X_0 in part (a).

(c) Let $(\mathbf{E},\mathbf{H}) \in D(M_{\mathrm{div}}^{(2)})$. We first observe that then $M(\mathbf{E},\mathbf{H}) \in X_{\mathrm{div}}^{(0)}$ by part (a) and that $M(\mathbf{E},\mathbf{H}) \in D(M^2)$. Moreover, $(\mathbf{E},\mathbf{H}) \in H^2(Q)^6$ by Proposition 7.11. To check that Ξ is contained in $H_{00}^1(Q)$ we differentiate (7.21) and obtain

$$\partial_j\Xi = \nabla\partial_j\sigma\cdot\mathbf{E} + \nabla\sigma\cdot\partial_j\mathbf{E} - \frac{\partial_j\sigma}{\varepsilon}\nabla\varepsilon\cdot\mathbf{E} + \frac{\sigma\partial_j\varepsilon}{\varepsilon^2}\nabla\varepsilon\cdot\mathbf{E} - \frac{\sigma}{\varepsilon}\nabla\partial_j\varepsilon\cdot\mathbf{E}$$
$$+ \frac{\sigma}{\varepsilon}\nabla\varepsilon\cdot\partial_j\mathbf{E} + \frac{\partial_j\sigma}{\varepsilon}\operatorname{div}(\varepsilon\mathbf{E}) - \frac{\sigma\partial_j\varepsilon}{\varepsilon^2}\operatorname{div}(\varepsilon\mathbf{E}) + \frac{\sigma}{\varepsilon}\partial_j\operatorname{div}(\varepsilon\mathbf{E}).$$

for all $j \in \{1,2,3\}$. The function $\partial_j\Xi$ thus belongs to $L^2(Q)$ due to the assumptions on ε and σ, the Sobolev embedding $H^1(Q) \hookrightarrow L^6(Q)$ and $\operatorname{div}(\varepsilon\mathbf{E}) \in H^1(Q)$. Using $\Xi \in L^2(Q)$

from part (b), we see that $\Xi = \operatorname{div}\big(\varepsilon(M(\mathbf{E},\mathbf{H}))_1\big)$ is an element of $H^1(Q)$. We observe that the map $f \mapsto \frac{\sigma}{\varepsilon} f$ belongs to $\mathcal{B}(L^2(\widehat{\Gamma}))$ and to $\mathcal{B}(H_0^1(\widehat{\Gamma}))$ and thus to $\mathcal{B}(H_0^{1/2}(\widehat{\Gamma}))$ by interpolation for each face $\widehat{\Gamma}$ of Q. This shows that $\frac{\sigma}{\varepsilon}\operatorname{div}(\varepsilon\mathbf{E})$ belongs to $H_0^{1/2}(\Gamma)$. The other terms on the right-hand side of (7.21) are contained in $W^{1,3}(Q)$ by Sobolev's embedding and the assumptions on ε and σ. Hence, they have traces in $W^{2/3,3}(\Gamma)$ by the Theorem 2.5.3 in [58]. By Proposition 7.11 and Lemma 7.5, the trace of

$$\varphi := (\partial_1\sigma)E_1 - \frac{\sigma}{\varepsilon}(\partial_1\varepsilon)E_1$$

vanishes on $\Gamma_2 \cup \Gamma_3$. As in the proof of Proposition 7.11 we construct smooth functions φ_n converging to φ in $W^{1,3}(Q)$ with support in the set $Q^{(1/n)}$, see (7.18). Their traces belong to $W_0^{2/3,3}(\Gamma_1)$ (which is the closure of $C_c^\infty(\Gamma_1)$ in $W^{2/3,3}(\Gamma_1)$) and converge in this space by Theorem 3.1 in [45]. Thus, $\operatorname{tr}(\varphi)$ in contained in $W_0^{2/3,3}(\Gamma_1)$ and its trace on $\partial\Gamma_1$ vanishes. Proposition 2.11, Remark 2.7 and Proposition 3.3 in [45] say that

$$H_0^{\theta,3}(\Gamma_1) = \big[L^3(\Gamma_1), W_0^{1,3}(\Gamma_1)\big]_\theta = \{\psi \in H^{\theta,3}(\Gamma_1) \mid \operatorname{tr}\psi = 0 \text{ on } \partial\Gamma_1\}$$

for $\theta > 0$, where $H_0^{\theta,3}(\Gamma_1)$ is the closure of the test functions in the Bessel potential space $H^{\theta,3}(\Gamma_1)$. Due to Proposition 1.4 and 1.3 in [52] we have for all $\theta \in (0, 2/3)$ the embedding

$$W^{2/3,3}(\Gamma_1) = (L^3(\Gamma_1), W^{1,3}(\Gamma_1))_{2/3,3} \hookrightarrow (L^3(\Gamma_1), W^{1,3}(\Gamma_1))_{\theta,1}$$

and the example on page 53 in [52] yields

$$(L^3(\Gamma_1), W^{1,3}(\Gamma_1))_{\theta,1} \hookrightarrow [L^3(\Gamma_1), W^{1,3}(\Gamma_1)]_\theta = H^{\theta,3}(\Gamma_1).$$

So, the space $W^{2/3,3}(\Gamma_1)$ is continuously embedded in $H^{\theta,3}(\Gamma_1)$ for any $\theta \in (0, 2/3)$. As a result, $\operatorname{tr}_{\Gamma_1}(\varphi)$ belongs to $H_0^{\theta,3}(\Gamma_1)$ for all $\theta \in (1/2, 2/3)$. Since $L^3(\Gamma_1) \hookrightarrow L^2(\Gamma_1)$ and $W_0^{1,3}(\Gamma_1) \hookrightarrow H_0^1(\Gamma_1)$, interpolation shows that $\operatorname{tr}_{\Gamma_1}(\varphi)$ is an element of

$$\big[L^2(\Gamma_1), H_0^1(\Gamma_1)\big]_\theta = \big(L^2(\Gamma_1), H_0^1(\Gamma_1)\big)_{\theta,2} \hookrightarrow H_0^{1/2}(\Gamma_1) = (L^2(\Gamma_1), H_0^1(\Gamma_1))_{1/2,2},$$

where we used see Corollary 4.37 in [52] for the first identity and Proposition 1.4 and 1.3 in [52] for the embedding. Summing up, φ is an element of $H_0^{1/2}(\Gamma)$. The remaining summands of Ξ can be treated similarly. Hence, Ξ belongs to $H_{00}^1(Q)$ and thus $M(\mathbf{E},\mathbf{H})$ to $X_{\mathrm{div}}^{(2)}$. □

Using this Proposition iteratively gives the following embedding and representations of the domains of the Maxwell operators.

Corollary 7.13. *We have the representations* $D\big((M_{\mathrm{div}}^{(0)})^j\big) = D(M^j) \cap X_{\mathrm{div}}^{(0)}$ *and, if* $\sigma = 0$, $D(M_0^j) = D(M^j) \cap X_0$, *for all* $j \in \mathbb{N}$. *Furthermore,* $X_{\mathrm{div}}^{(2)}$ *is continuously embedded into* $D\big((M_{\mathrm{div}}^{(0)})^2\big)$.

7.2.2. The splitting operators and their domains

The basic idea for the splitting scheme proposed in [75] is to split the curl operator into

$$\operatorname{curl} = C_1 - C_2$$

with

$$C_1 := \begin{pmatrix} 0 & 0 & \partial_2 \\ \partial_3 & 0 & 0 \\ 0 & \partial_1 & 0 \end{pmatrix} \qquad \text{and} \qquad C_2 := \begin{pmatrix} 0 & \partial_3 & 0 \\ 0 & 0 & \partial_1 \\ \partial_2 & 0 & 0 \end{pmatrix} \tag{7.22}$$

and to define the *splitting operators*

$$A := \begin{pmatrix} -\frac{\sigma}{2\varepsilon} I & \frac{1}{\varepsilon} C_1 \\ \frac{1}{\mu} C_2 & 0 \end{pmatrix} \qquad \text{and} \qquad B := \begin{pmatrix} -\frac{\sigma}{2\varepsilon} I & -\frac{1}{\varepsilon} C_2 \\ -\frac{1}{\mu} C_1 & 0 \end{pmatrix}. \tag{7.23}$$

These operators are endowed with the domains

$$\begin{aligned} D(A) := \{(u,v) \in X \mid (C_1 v, C_2 u) \in X, \\ u_1 = 0 \text{ on } \Gamma_2, \;\; u_2 = 0 \text{ on } \Gamma_3, \;\; u_3 = 0 \text{ on } \Gamma_1\}, \\ D(B) := \{(u,v) \in X \mid (C_2 v, C_1 u) \in X, \\ u_1 = 0 \text{ on } \Gamma_3, \;\; u_2 = 0 \text{ on } \Gamma_1, \;\; u_3 = 0 \text{ on } \Gamma_2\}, \end{aligned}$$

which contain "partial" Dirichlet boundary conditions. Observe that the boundary conditions of M have been partitioned into the boundary conditions of the operators A and B. This is done in such a way that the square integrability of the corresponding derivatives assures that the boundary conditions are well-defined, see Theorem 4.12 in [1]. Clearly, we have

$$D(A) \cap D(B) \hookrightarrow D(M) \qquad \text{and} \qquad M = A + B \quad \text{on } D(A) \cap D(B).$$

Keep in mind that neither the divergence conditions nor the boundary condition for the magnetic field have been taken into account in the definition of A and B. We write A_0 and B_0 for the operator A, respectively B, with $\sigma = 0$, i.e. we have $D(A_0) = D(A)$, $D(B_0) = D(B)$,

$$A_0 = A + \begin{pmatrix} \frac{\sigma}{2\varepsilon} I & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{and} \qquad B_0 = B + \begin{pmatrix} \frac{\sigma}{2\varepsilon} I & 0 \\ 0 & 0 \end{pmatrix}. \tag{7.24}$$

The following statements can be found in Section 4.3 in [37]. Let $u, \psi \in L^2(Q)^3$ with $C_1\psi \in L^2(Q)^3$ and $C_2 u \in L^2(Q)^3$. Let furthermore the boundary conditions

$$\begin{aligned} u_2 = 0 \quad \text{or} \quad \psi_1 = 0 \qquad \text{on} \quad \Gamma_3, \\ u_3 = 0 \quad \text{or} \quad \psi_2 = 0 \qquad \text{on} \quad \Gamma_1, \\ u_1 = 0 \quad \text{or} \quad \psi_3 = 0 \qquad \text{on} \quad \Gamma_2, \end{aligned} \tag{7.25}$$

hold true. Then we see with integration by parts that

$$(C_2 u \mid \psi)_{L^2} = (u \mid -C_1 \psi)_{L^2}. \tag{7.26}$$

Let $v, \varphi \in L^2(Q)^3$ with $C_1 v \in L^2(Q)^2$ and $C_2 \varphi \in L^2(Q)^3$. Let additionally the boundary conditions

$$\begin{aligned} v_3 = 0 \quad \text{or} \quad \varphi_1 = 0 \qquad \text{on} \quad \Gamma_2, \\ v_1 = 0 \quad \text{or} \quad \varphi_2 = 0 \qquad \text{on} \quad \Gamma_3, \\ v_2 = 0 \quad \text{or} \quad \varphi_3 = 0 \qquad \text{on} \quad \Gamma_1, \end{aligned} \tag{7.27}$$

be satisfied. Then we get with integration by parts that

$$(C_1 v \mid \varphi)_{L^2} = (v \mid -C_2 \varphi)_{L^2}. \tag{7.28}$$

This gives us the adjoint operators of A, B and M.

Lemma 7.14. *(a) The adjoints of the splitting operators have the domains*

$$D(A^*) = D(A_0^*) = D(A) \qquad \text{and} \qquad D(B^*) = D(B_0^*) = D(B)$$

and satisfy the identities $A_0^* = -A_0$, $B_0^* = -B_0$,

$$A^* = \begin{pmatrix} -\frac{\sigma}{2\varepsilon} I & -\frac{1}{\varepsilon} C_1 \\ -\frac{1}{\mu} C_2 & 0 \end{pmatrix} \qquad \text{and} \qquad B^* = \begin{pmatrix} -\frac{\sigma}{2\varepsilon} I & \frac{1}{\varepsilon} C_2 \\ \frac{1}{\mu} C_1 & 0 \end{pmatrix}.$$

(b) The adjoint of M is given by $D(M^*) = D(M)$ and

$$M^* = A^* + B^* = \begin{pmatrix} -\frac{\sigma}{\varepsilon} I & -\frac{1}{\varepsilon} \operatorname{curl} \\ \frac{1}{\mu} \operatorname{curl} & 0 \end{pmatrix}.$$

Proof:
(a) The domains of and the formulas for the operators A_0^* and B_0^* follow from (7.25), (7.27), (7.26) and (7.28), see Lemma 4.3 in [37]. Together with the symmetry and boundedness of $\Sigma = \begin{pmatrix} -\frac{\sigma}{2\varepsilon} I & 0 \\ 0 & 0 \end{pmatrix}$ we thus obtain

$$A^* = A_0^* + \Sigma^* = -A_0 + \Sigma = \begin{pmatrix} -\frac{\sigma}{2\varepsilon} I & -\frac{1}{\varepsilon} C_1 \\ -\frac{1}{\mu} C_2 & 0 \end{pmatrix}$$

on $D(A^*) = D(A_0^*)$. Analogously, we see $D(B^*) = D(B_0^*)$ and $B^* = \begin{pmatrix} -\frac{\sigma}{2\varepsilon} I & \frac{1}{\varepsilon} C_2 \\ \frac{1}{\mu} C_1 & 0 \end{pmatrix}$.
(b) The skew-adjointness of M with $\sigma = 0$ was shown for instance in Proposition 3.5 in [37]. One can then proceed as above. □

As usual, we set

$$D(AB) := \{u \in D(B) \mid Bu \in D(A)\}$$

and analougously for $D(A^2)$, $D(BA)$ and $D(B^2)$. Further properties of the splitting operators are shown in the Sections 8.1, 8.2 and 8.3.

The domains of the Maxwell operators are embedded into some domains of the splitting operators.

Proposition 7.15. *(a) $D(M^{(0)}_{\mathrm{div}})$ is continuously embedded into $D(A)$ and $D(B)$. Moreover, we have*

$$\|A(u,v)\|_X \le c\big(\|(u,v)\|_{X^{(0)}_{\mathrm{div}}} + \|M(u,v)\|_{X^{(0)}_{\mathrm{div}}}\big),$$
$$\|B(u,v)\|_X \le c\big(\|(u,v)\|_{X^{(0)}_{\mathrm{div}}} + \|M(u,v)\|_{X^{(0)}_{\mathrm{div}}}\big)$$

for all $(u,v) \in D(M^{(0)}_{\mathrm{div}})$, with the constants depending only on $\|\varepsilon\|_{W^{1,\infty}}$, $\|\mu\|_{W^{1,\infty}}$, $\|\sigma\|_{L^\infty}$ and δ.

(b) Let $\varepsilon, \mu \in W^{2,3}(Q)$. Then $X^{(2)}_{\mathrm{div}}$ is continuously embedded into $D(A^2)$, $D(AB)$, $D(BA)$ and $D(B^2)$. Furthermore, we have

$$\big\|A^2(u,v)\big\|_X \le c\,\|(u,v)\|_{X^{(2)}_{\mathrm{div}}},$$
$$\|AB(u,v)\|_X \le c\,\|(u,v)\|_{X^{(2)}_{\mathrm{div}}},$$
$$\|BA(u,v)\|_X \le c\,\|(u,v)\|_{X^{(2)}_{\mathrm{div}}},$$
$$\big\|B^2(u,v)\big\|_X \le c\,\|(u,v)\|_{X^{(2)}_{\mathrm{div}}}$$

for all $(u,v) \in X^{(2)}_{\mathrm{div}}$, with the constants c only depending on $\|\varepsilon\|_{W^{1,\infty}\cap W^{2,3}}$, $\|\mu\|_{W^{1,\infty}\cap W^{2,3}}$, $\|\sigma\|_{W^{1,\infty}}$ and δ.

(c) We have $M^{(0)}_{\mathrm{div}} = A + B$ on $D(M^{(0)}_{\mathrm{div}})$ and $M^{(2)}_{\mathrm{div}} = A + B$ on $D(M^{(2)}_{\mathrm{div}})$.

Proof:
(a) Let $(u,v) \in D(M^{(0)}_{\mathrm{div}})$. Then (u,v) satisfies the boundary conditions of $D(A)$ and $D(B)$ due to Proposition 7.11. The embedding $D(M^{(1)}_{\mathrm{div}}) \hookrightarrow H^1(Q)^6$ from Proposition 7.11 then implies that (u,v) is contained in $D(A) \cap D(B)$. The embedding follows from the obvious estimate

$$\max\Big\{\|A(u,v)\|_{L^2}, \|B(u,v)\|_{L^2}\Big\} \le c\,\|(u,v)\|_{H^1}$$

and the inequality $\|(u,v)\|_{H^1} \le c\,\|(u,v)\|_{D(M^{(0)}_{\mathrm{div}})}$ in Proposition 7.11. Here the constants only depend on the claimed quantities.

(b) Let $(u,v) \in X^{(2)}_{\mathrm{div}}$. For the first component of $A(u,v)$ and $B(u,v)$ we have the traces

$$-\frac{\sigma}{2\varepsilon}u_1 + \frac{1}{\varepsilon}\partial_2 v_3 = 0 \quad \text{on } \Gamma_3 \qquad \text{and} \qquad -\frac{\sigma}{2\varepsilon}u_1 - \frac{1}{\varepsilon}\partial_3 v_2 = 0 \quad \text{on } \Gamma_2,$$

respectively, due to Proposition 7.11 and Lemma 7.5. Thus, $A(u,v)$ fulfils the boundary condition of $D(B)$ and $B(u,v)$ fulfils the boundary condition of $D(A)$. Taking additionally $M(u,v) \in D(M)$ into account, we obtain

$$\begin{aligned} -\frac{\sigma}{2\varepsilon}u_1 + \frac{1}{\varepsilon}\partial_2 v_3 = (M(u,v))_{1,1} + \frac{\sigma}{2\varepsilon}u_1 + \frac{1}{\varepsilon}\partial_3 v_2 = 0 \quad \text{on } \Gamma_2, \\ -\frac{\sigma}{2\varepsilon}u_1 - \frac{1}{\varepsilon}\partial_3 v_2 = (M(u,v))_{1,1} + \frac{\sigma}{2\varepsilon}u_1 - \frac{1}{\varepsilon}\partial_2 v_3 = 0 \quad \text{on } \Gamma_3, \end{aligned}$$

respectively, where $(M(u,v))_{1,1}$ denotes the first component of $(M(u,v))_1$. Thus, the boundary conditions of $D(A)$ is satisfied by $A(u,v)$ and the boundary condition of $D(B)$ is satisfied by $B(u,v)$. The second and third component of $A(u,v)$ and $B(u,v)$ are treated similarly. Together with the embedding $X^{(2)}_{\mathrm{div}} \hookrightarrow H^2(Q)^6$ from Proposition 7.11 we have shown that (u,v) is contained in $D(A^2) \cap D(AB) \cap D(BA) \cap D(B^2)$. The continuity of the embedding follows from the estimate

$$\max\Big\{ \big\|A^2(u,v)\big\|_{L^2}, \|AB(u,v)\|_{L^2}, \|BA(u,v)\|_{L^2}, \big\|B^2(u,v)\big\|_{L^2} \Big\} \le c\,\|(u,v)\|_{H^2}$$

and the estimate $\|(u,v)\|_{H^2} \le c\,\|(u,v)\|_{X^{(2)}_{\mathrm{div}}}$ from Proposition 7.11, with the constants only depending on the claimed quantities.

The statement of part (c) is clear. □

For our error analysis we need versions of the splitting operators in an H^1-setting and in an H^2-setting. For the splitting in the H^1-setting, we define the space

$$Y := \{(u,v) \in H^1(Q)^6 \mid u_j = 0 \text{ on } \Gamma \setminus \Gamma_j, \;\; v_j = 0 \text{ on } \Gamma_j, \text{ for all } j \in \{1,2,3\}\}. \tag{7.29}$$

We use for $(u,v),(\varphi,\psi) \in Y$ the weighted inner product

$$\big((u,v) \mid (\varphi,\psi)\big)_Y := \int_Q \Big(\varepsilon u \cdot \varphi + \mu v \cdot \psi + \varepsilon \sum_{j=1}^{3} \partial_j u \cdot \partial_j \varphi + \mu \sum_{j=1}^{3} \partial_j v \cdot \partial_j \psi \Big)\, \mathrm{d}x$$

with the induced norm $\|\cdot\|_Y$. Due to our assumptions on ε and μ, this norm is equivalent to the H^1-norm. The continuity of the traces implies that Y is a closed subspace of $H^1(Q)^6$.

Remark 7.16. *By definition we have*

$$Y \hookrightarrow D(A) \cap D(B) \cap D(A^*) \cap D(B^*).$$

The part of A in Y is the operator A_Y with domain

$$D(A_Y) := \{(u,v) \in Y \mid (u,v) \in D(A),\; A(u,v) \in Y\}$$

and $A_Y(u,v) := A(u,v)$ for $(u,v) \in D(A_Y)$, and the part of B in Y is the operator B_Y with

$$D(B_Y) := \{(u,v) \in Y \mid (u,v) \in D(B),\ B(u,v) \in Y\}$$

and $B_Y(u,v) := B(u,v)$ for $(u,v) \in D(B_Y)$. Combining the formulas for A and B with the definition of Y and the assumptions on ε, μ and σ yields, due to Lemma 7.2, the following representation for $D(A_Y)$ and $D(B_Y)$. It will be improved in Corollary 8.3.

Lemma 7.17. *We have*

$$\begin{aligned} D(A_Y) &= \{(u,v) \in Y \mid (C_1 v, C_2 u) \in Y\} \\ &= \{(u,v) \in H^1(Q)^6 \mid u_j = 0 \text{ on } \Gamma \setminus \Gamma_j,\ \ v_j = 0 \text{ on } \Gamma_j, \\ &\qquad \text{for all } j \in \{1,2,3\}, \\ &\qquad \partial_2 u_1, \partial_3 u_2, \partial_1 u_3, \partial_3 v_1, \partial_1 v_2, \partial_2 v_3 \in H^1(Q), \\ &\qquad \partial_3 v_1 = 0 \text{ on } \Gamma \setminus \Gamma_2,\ \ \partial_1 v_2 = 0 \text{ on } \Gamma \setminus \Gamma_3,\ \ \partial_2 v_3 = 0 \text{ on } \Gamma \setminus \Gamma_1, \\ &\qquad \partial_3 u_2 = 0 \text{ on } \Gamma_1,\ \ \partial_1 u_3 = 0 \text{ on } \Gamma_2,\ \ \partial_2 u_1 = 0 \text{ on } \Gamma_3\} \end{aligned}$$

and

$$\begin{aligned} D(B_Y) &= \{(u,v) \in Y \mid (C_2 v, C_1 u) \in Y\} \\ &= \{(u,v) \in H^1(Q)^6 \mid u_j = 0 \text{ on } \Gamma \setminus \Gamma_j,\ v_j = 0 \text{ on } \Gamma_j, \\ &\qquad \text{for all } j \in \{1,2,3\}, \\ &\qquad \partial_3 u_1, \partial_1 u_2, \partial_2 u_3, \partial_2 v_1, \partial_3 v_2, \partial_1 v_3 \in H^1(Q), \\ &\qquad \partial_2 v_1 = 0 \text{ on } \Gamma \setminus \Gamma_3,\ \partial_3 v_2 = 0 \text{ on } \Gamma \setminus \Gamma_1,\ \ \partial_1 v_3 = 0 \text{ on } \Gamma \setminus \Gamma_2, \\ &\qquad \partial_3 u_1 = 0 \text{ on } \Gamma_2,\ \ \partial_1 u_2 = 0 \text{ on } \Gamma_3,\ \ \partial_2 u_3 = 0 \text{ on } \Gamma_1\}. \end{aligned}$$

Proof:
The identity

$$D(A_Y) = \{(u,v) \in Y \mid (C_1 v, C_2 u) \in Y\}$$

follows from $Y \hookrightarrow D(A)$, the general assumptions on ε, μ and σ, and Lemma 7.2. The second equality in the reformulation of $D(A_Y)$ is true by the definition of Y. The operator B_Y is treated in the same way. □

For the splitting in the H^2-setting, we define the space

$$\begin{aligned} Z := \{(u,v) \in H^2(Q)^6 \mid u_j = 0 \text{ on } \Gamma \setminus \Gamma_j,\ \ v_j = 0 \text{ on } \Gamma_j, \text{¸notag} \qquad (7.30) \\ \partial_j u_j = 0 \text{ on } \Gamma_j, \text{ for all } j \in \{1,2,3\}, \qquad (7.31) \\ \partial_j v_k = 0 \text{ on } \Gamma_j \text{ for all } j,k \in \{1,2,3\} \text{ with } j \neq k\} \end{aligned}$$

and use for $(u,v),(\varphi,\psi) \in Z$ the weighted inner product

$$\big((u,v) \mid (\varphi,\psi)\big)_Z := \int_Q \Big(\varepsilon u\cdot\varphi + \mu v\cdot\psi + \varepsilon\sum_{j=1}^{3}\partial_j u\cdot\partial_j\varphi + \mu\sum_{j=1}^{3}\partial_j v\cdot\partial_j\psi \\ + \varepsilon\sum_{j,k=1}^{3}\partial_{jk}u\cdot\partial_{jk}\varphi + \mu\sum_{j,k=1}^{3}\partial_{jk}v\cdot\partial_{jk}\psi\Big)\,\mathrm{d}x.$$

Due to the general assumptions on ε and μ, the norm $\|\cdot\|_Z$ that is induced by this inner product is equivalent to the H^2-norm. We have $Z \hookrightarrow D(A)\cap D(B) \subseteq D(M)$ and by the continuity of the traces that $Z \subseteq H^2(Q)^6$ is closed. We define the restriction A_Z of A to the subspace

$$D(A_Z) := \{(u,v)\in Z \mid \partial_2 u_1, \partial_3 u_2, \partial_1 u_3, \partial_3 v_1, \partial_1 v_2, \partial_2 v_3 \in H^2(Q), \\ \partial_{22}u_1 = 0 \text{ on } \Gamma_2,\ \ \partial_{33}u_2 = 0 \text{ on } \Gamma_3,\ \ \partial_{11}u_3 = 0 \text{ on } \Gamma_1\} \tag{7.32}$$

of Z by $A_Z(u,v) := A(u,v)$ for $(u,v)\in D(A_Z)$ and the restriction B_Z of B to the subspace

$$D(B_Z) := \{(u,v)\in Z \mid \partial_3 u_1, \partial_1 u_2, \partial_2 u_3, \partial_2 v_1, \partial_3 v_2, \partial_1 v_3 \in H^2(Q), \\ \partial_{33}u_1 = 0 \text{ on } \Gamma_3,\ \ \partial_{11}u_2 = 0 \text{ on } \Gamma_1,\ \ \partial_{22}u_3 = 0 \text{ on } \Gamma_2\} \tag{7.33}$$

of Z by $B_Z(u,v) := B(u,v)$ for $(u,v)\in D(B_Z)$. Note that in contrast to the analogous H^1-setting, A_Z and B_Z are not the parts of A and B in Z, respectively. This change is necessary due to some technical difficulties in later proofs. We now enforce that A_Z and B_Z map into Z by posing a trace condition on the coefficients.

Lemma 7.18. *If $\varepsilon,\mu,\sigma \in W^{2,3}(Q)$ and $\partial_\nu\varepsilon = \partial_\nu\mu = \partial_\nu\sigma = 0$ on Γ, then A_Z and B_Z map into Z.*

Proof:
Let $(u,v)\in D(A_Z)$. The smoothness $A(u,v)\in H^2(Q)^6$ follows from the assumptions on ε, μ and σ and Lemma 7.2. In the rest of this proof we use Lemma 7.10 and 7.5 frequently and without further mentioning. The first component of $A(u,v)$ satisfies the zero-order boundary conditions $-\frac{\sigma}{2\varepsilon}u_1 + \frac{1}{\varepsilon}\partial_2 v_3 = 0$ on $\Gamma_2\cup\Gamma_3$ due to the definition of Z. We further obtain

$$\partial_1\big(\tfrac{1}{\varepsilon}\partial_2 v_3\big) = -\frac{\partial_1\varepsilon}{\varepsilon^2}\partial_2 v_3 + \frac{1}{\varepsilon}\partial_2\partial_1 v_3 = 0$$

on Γ_1 due to $\partial_\nu\varepsilon = 0$ on Γ and $\partial_1 v_3 = 0$ on Γ_1 from the definition of Z. Moreover, the equation

$$\partial_1\big(-\tfrac{\sigma}{2\varepsilon}u_1\big) = -\frac{\partial_1\sigma}{2\varepsilon}u_1 + \frac{\sigma\partial_1\varepsilon}{2\varepsilon^2}u_1 - \frac{\sigma}{2\varepsilon}\partial_1 u_1 = 0$$

on Γ_1 follows from $\partial_\nu\sigma = \partial_\nu\varepsilon = 0$ on Γ and $\partial_1 u_1 = 0$ on Γ_1 from the definition of Z. The first component of $A(u,v)$ then fulfils the boundary conditions of Z.

The zero-order boundary condition $\frac{1}{\mu}\partial_3 u_2 = 0$ on Γ_1 of the forth component of $A(u,v)$ is satisfied by the definition of Z. Using $\partial_2 u_2 = 0$ on Γ_2 by the definition of Z and $\partial_\nu \mu = 0$ on Γ, we infer

$$\partial_2\big(\tfrac{1}{\mu}\partial_3 u_2\big) = -\frac{\partial_2\mu}{\mu^2}\partial_3 u_2 + \frac{1}{\mu}\partial_3\partial_2 u_2 = 0$$

on Γ_2. Again due to $\partial_\nu\mu = 0$ on Γ and this time due to the definition of $D(A_Z)$ we compute

$$\partial_3\big(\tfrac{1}{\mu}\partial_3 u_2\big) = -\frac{\partial_3\mu}{\mu^2}\partial_3 u_2 + \frac{1}{\mu}\partial_{33} u_2 = 0$$

on Γ_3. Hence, the boundary conditions of Z of the fourth component of $A(u,v)$ are shown. The other components of $A(u,v)$ are treated analogously.

Let $(u,v) \in D(B_Z)$. In the same way as for $D(A_Z)$ we check $B(u,v) \in H^2(Q)^6$, the boundary conditions

$$-\frac{\sigma}{2\varepsilon}u_1 - \frac{1}{\varepsilon}\partial_3 v_2 = 0 \qquad \text{on } \Gamma_2 \cup \Gamma_3$$

and

$$\partial_1\big(-\tfrac{\sigma}{2\varepsilon}u_1 - \tfrac{1}{\varepsilon}\partial_3 v_2\big) = -\frac{\partial_1\sigma}{2\varepsilon}u_1 + \frac{\sigma\partial_1\varepsilon}{2\varepsilon^2}u_1 - \frac{\sigma}{2\varepsilon}\partial_1 u_1 + \frac{\partial_1\varepsilon}{\varepsilon^2}\partial_3 v_2 - \frac{1}{\varepsilon}\partial_3\partial_1 v_2 = 0 \qquad \text{on } \Gamma_1$$

of the first component of $B(u,v)$, as well as the boundary conditions

$$\begin{aligned}
-\frac{1}{\mu}\partial_2 u_3 = 0 \qquad & \text{on } \Gamma_1,\\
\partial_2\big(-\tfrac{1}{\mu}\partial_2 u_3\big) = \frac{\partial_2\mu}{\mu^2}\partial_2 u_3 - \frac{1}{\mu}\partial_{22} u_3 = 0 \qquad & \text{on } \Gamma_2,\\
\partial_3\big(-\tfrac{1}{\mu}\partial_2 u_3\big) = \frac{\partial_3\mu}{\mu^2}\partial_2 u_3 - \frac{1}{\mu}\partial_2\partial_3 u_3 = 0 \qquad & \text{on } \Gamma_3
\end{aligned}$$

of the forth component of $B(u,v)$. The other components of $B(u,v)$ are treated analogously. □

7.3. Solutions to the Maxwell equations

Observe that the electric boundary condition has been built into the domain of the Maxwell operator. The divergence condition on the magnetic field and the magnetic boundary condition are conserved quantities, see Chapter 1 in [56]. Rather than at (7.1), we thus look at the inhomogeneous Cauchy problem

$$\partial_t \begin{pmatrix}\mathbf{E}(t)\\ \mathbf{H}(t)\end{pmatrix} = M\begin{pmatrix}\mathbf{E}(t)\\ \mathbf{H}(t)\end{pmatrix} + \begin{pmatrix}-\frac{1}{\varepsilon}\mathbf{J}_0(t)\\ 0\end{pmatrix} \qquad \text{in } Q, \tag{7.34a}$$

$$(\mathbf{E}(0),\mathbf{H}(0)) = (\mathbf{E}_0,\mathbf{H}_0) \in D(M), \tag{7.34b}$$

with $\rho(t) := \operatorname{div}(\varepsilon \mathbf{E}(t))$ in $L^2(Q)$ or $H^1(Q)$, assuming that $\operatorname{div}(\varepsilon \mathbf{E}_0)$ belongs to $L^2(Q)$ or $H^1(Q)$, respectively, and that

$$\operatorname{div}(\mu \mathbf{H}_0) = 0 \quad \text{on} \quad Q \qquad \text{and} \qquad \operatorname{tr}_n(\mu \mathbf{H}_0) = 0 \quad \text{on} \quad \Gamma.$$

We look for solutions $(\mathbf{E}, \mathbf{H})$ that (at least) belong to $C^1([0,\infty), X) \cap C([0,\infty), D(M))$.

First, we look at problem (7.1) without the divergence conditions and without the magnetic boundary condition, i.e.

$$\partial_t \mathbf{E}(t) = \frac{1}{\varepsilon} \operatorname{curl} \mathbf{H}(t) - \frac{1}{\varepsilon}\big(\sigma \mathbf{E}(t) + \mathbf{J}_0(t)\big) \qquad \text{in } Q, \tag{7.35a}$$

$$\partial_t \mathbf{H}(t) = -\frac{1}{\mu} \operatorname{curl} \mathbf{E}(t) \qquad \text{in } Q, \tag{7.35b}$$

$$\operatorname{tr}_t(\mathbf{E}(t)) = 0 \qquad \text{on } \Gamma, \tag{7.35c}$$

$$\mathbf{E}(0) = \mathbf{E}_0, \qquad \mathbf{H}(0) = \mathbf{H}_0 \qquad \text{in } Q, \tag{7.35d}$$

for $t \geq 0$.

Proposition 7.19. *(a) The operator M generates a contraction C_0-semigroup e^{tM} on X. If $(\boldsymbol{E}_0, \boldsymbol{H}_0) \in D(M)$ and $\big(\boldsymbol{J}_0, 0\big) \in C([0,\infty), D(M)) + C^1([0,\infty), X)$, then there exists a unique solution $(\boldsymbol{E}, \boldsymbol{H}) \in C^1([0,\infty), X) \cap C([0,\infty), D(M))$ to (7.35), which fulfils*

$$(\boldsymbol{E}(t), \boldsymbol{H}(t)) = e^{tM}(\boldsymbol{E}_0, \boldsymbol{H}_0) - \int_0^t e^{(t-s)M}\big(\tfrac{1}{\varepsilon}\boldsymbol{J}_0(s), 0\big)\,\mathrm{d}s \qquad \text{in } L^2(Q)^6, \tag{7.36a}$$

$$\operatorname{div}(\varepsilon \boldsymbol{E}(t)) = e^{-\frac{\sigma}{\varepsilon}t} \operatorname{div}(\varepsilon \boldsymbol{E}_0) - \int_0^t e^{-\frac{\sigma}{\varepsilon}(t-s)}\Big(\nabla\big(\tfrac{\sigma}{\varepsilon}\big)\varepsilon \boldsymbol{E}(s) + \operatorname{div}(\boldsymbol{J}_0(s))\Big)\,\mathrm{d}s \tag{7.36b}$$

$$\text{in } H^{-1}(Q),$$

$$\operatorname{div}(\varepsilon \boldsymbol{E}(t)) = \operatorname{div}(\varepsilon \boldsymbol{E}_0) - \int_0^t \operatorname{div}(\sigma \boldsymbol{E}(s) + \boldsymbol{J}_0(s))\,\mathrm{d}s \qquad \text{in } H^{-1}(Q), \tag{7.36c}$$

$$\operatorname{div}(\mu \boldsymbol{H}(t)) = \operatorname{div}(\mu \boldsymbol{H}_0) \qquad \text{in } H^{-1}(Q), \tag{7.36d}$$

$$\operatorname{tr}_n(\mu \boldsymbol{H}(t)) = \operatorname{tr}_n(\mu \boldsymbol{H}_0) \qquad \text{in } H^{-1/2}(\Gamma), \tag{7.36e}$$

for all $t \geq 0$. If $\sigma = 0$, then the semigroup can be extended to a unitary group.

(b) Let $(\boldsymbol{E}_0, \boldsymbol{H}_0) \in X$ and $(\boldsymbol{J}_0, 0) \in L^1_{loc}([0,\infty), X)$. Define $(\boldsymbol{E}(t), \boldsymbol{H}(t))$ by (7.36a). Then the equations (7.36b), (7.36c) and (7.36d) still hold true in $H^{-1}(Q)$.

Proof:
(a) We define the operator $(\widetilde{M}, D(\widetilde{M}))$ on X with $D(\widetilde{M}) := D(M)$ and

$$\widetilde{M} := \begin{pmatrix} 0 & \frac{1}{\varepsilon}\operatorname{curl} \\ -\frac{1}{\mu}\operatorname{curl} & 0 \end{pmatrix}.$$

Let $(\mathbf{E}_0, \mathbf{H}_0) \in D(M) = D(\widetilde{M})$. Due to Proposition 3.5 in [37] the operator $\widetilde{M}$ generates a unitary C_0-group $e^{t\widetilde{M}}$ and $e^{t\widetilde{M}}(\mathbf{E}_0, \mathbf{H}_0)$ is the unique solution to (7.35) in $C^1([0,\infty), X) \cap C([0,\infty), D(M))$ if $\sigma = 0$ and $\mathbf{J}_0 = 0$. Because $M - \widetilde{M}$ is bounded and dissipative on X, Theorem III.2.7 in [23] yields that M generates a contractive C_0-semigroup e^{tM} on X. Under the assumptions on $(\mathbf{E}_0, \mathbf{H}_0)$ and $(\mathbf{J}_0, 0)$ we thus obtain a solution

$$(\mathbf{E}, \mathbf{H}) \in C^1([0,\infty), X) \cap C([0,\infty), D(M))$$

to (7.35), given by (7.36a).

Equation (7.35a) implies

$$\begin{aligned}\partial_s \operatorname{div}(\varepsilon \mathbf{E}(s)) &= \operatorname{div}\operatorname{curl}(\varepsilon \mathbf{E}(s)) - \operatorname{div}\big(\tfrac{\sigma}{\varepsilon}\varepsilon \mathbf{E}(s)\big) - \operatorname{div}(\mathbf{J}_0(s)) \\ &= -\frac{\sigma}{\varepsilon}\operatorname{div}(\varepsilon \mathbf{E}(s)) - \nabla\big(\tfrac{\sigma}{\varepsilon}\big)\varepsilon \mathbf{E}(s) - \operatorname{div}(\mathbf{J}_0(s)),\end{aligned}$$

so that

$$\begin{aligned}\partial_s\big(e^{\frac{\sigma}{\varepsilon}s}\operatorname{div}(\varepsilon \mathbf{E}(s)\big) &= \frac{\sigma}{\varepsilon}e^{\frac{\sigma}{\varepsilon}s}\operatorname{div}(\varepsilon \mathbf{E}(s)) - e^{\frac{\sigma}{\varepsilon}s}\big(\tfrac{\sigma}{\varepsilon}\operatorname{div}(\varepsilon \mathbf{E}(s)) + \nabla\big(\tfrac{\sigma}{\varepsilon}\big)\varepsilon \mathbf{E}(s) + \operatorname{div}(\mathbf{J}_0(s))\big) \\ &= -e^{\frac{\sigma}{\varepsilon}s}\Big(\nabla\big(\tfrac{\sigma}{\varepsilon}\big)\varepsilon \mathbf{E}(s) + \operatorname{div}(\mathbf{J}_0(s))\Big)\end{aligned}$$

in $H^{-1}(Q)$ for $s \geq 0$. Integration from 0 to t yields

$$e^{\frac{\sigma}{\varepsilon}t}\operatorname{div}(\varepsilon \mathbf{E}(t)) = \operatorname{div}(\varepsilon \mathbf{E}_0) - \int_0^t e^{\frac{\sigma}{\varepsilon}s}\Big(\nabla\big(\tfrac{\sigma}{\varepsilon}\big)\varepsilon \mathbf{E}(s) + \operatorname{div}(\mathbf{J}_0(s))\Big)\,\mathrm{d}s$$

and thus

$$\operatorname{div}(\varepsilon \mathbf{E}(t)) = e^{-\frac{\sigma}{\varepsilon}t}\operatorname{div}(\varepsilon \mathbf{E}_0) - \int_0^t e^{-\frac{\sigma}{\varepsilon}(t-s)}\Big(\nabla\big(\tfrac{\sigma}{\varepsilon}\big)\varepsilon \mathbf{E}(s) + \operatorname{div}(\mathbf{J}_0(s))\Big)\,\mathrm{d}s$$

in $H^{-1}(Q)$, which is (7.36b).

Let $\varphi \in H_0^1(Q)$. Again equation (7.35a) and $\operatorname{div}\operatorname{curl} = 0$ in $H^{-1}(Q)$ yield the formula

$$\begin{aligned}\partial_t \langle \operatorname{div}(\varepsilon \mathbf{E}(t)), \varphi\rangle_{H^{-1}(Q),H_0^1(Q)} &= -\partial_t \int_Q \varepsilon \mathbf{E}(t)\cdot\nabla\varphi\,\mathrm{d}x \\ &= -\int_Q (\operatorname{curl}\mathbf{H}(t) - \sigma\mathbf{E}(t) - \mathbf{J}_0(t))\cdot\nabla\varphi\,\mathrm{d}x \\ &= \langle -\operatorname{div}(\sigma\mathbf{E}(t) + \mathbf{J}_0(t)), \varphi\rangle_{H^{-1}(Q),H_0^1(Q)}\end{aligned}$$

since $(\mathbf{J}_0, 0) \in C([0,\infty), D(M)) \cap C^1([0,\infty), X)$. Hence,

$$\partial_t \operatorname{div}(\varepsilon \mathbf{E}(t)) = -\operatorname{div}(\sigma\mathbf{E}(t) + \mathbf{J}_0(t))$$

in $H^{-1}(Q)$. By integrating, (7.36c) is thus valid in $H^{-1}(Q)$. In the same way one shows (7.36d) by means of (7.35b).

To derive (7.36e), we take $\varphi \in H^2(Q)$. Again (7.35b) and Proposition 7.3 imply that

$$\begin{aligned} 0 &= \big(\partial_t(\mu \mathbf{H}(t)) + \operatorname{curl} \mathbf{E}(t) \mid \nabla\varphi\big)_{L^2} \\ &= -\partial_t \int_Q \operatorname{div}(\mu\mathbf{H}(t))\varphi \,\mathrm{d}x + \partial_t \Big\langle \operatorname{tr}_n(\mu\mathbf{H}(t)), \varphi \Big\rangle_{H^{-1/2}(\Gamma), H^{1/2}(\Gamma)} \\ &\quad + \int_Q \mathbf{E}(t) \cdot \operatorname{curl} \nabla\varphi \,\mathrm{d}x - \Big\langle \operatorname{tr}_t(\mathbf{E}(t)), \nabla\varphi \Big\rangle_{H^{-1/2}(\Gamma)^3, H^{1/2}(\Gamma)^3}. \end{aligned}$$

Using (7.36d), $\operatorname{curl}\nabla = 0$ and (7.35c), we deduce that

$$\partial_t \left\langle \operatorname{tr}_n(\mu\mathbf{H}(t)), \varphi \right\rangle_{H^{-1/2}(\Gamma), H^{1/2}(\Gamma)} = 0.$$

This implies (7.36e) since $H^2(Q)$ is dense in $H^1(Q)$ and the trace map $\operatorname{tr} : H^1(Q) \to H^{1/2}(\Gamma)$ is surjective by Proposition 7.3.

For $\sigma = 0$ the semigroup can be extended to a unitary group by Stone's Theorem since M is skew-adjoint by Proposition 3.5 in [37].

The statement of (b) is seen by approximation. □

Before we can continue with the generation properties of the restricted Maxwell operators, we show a weaker version of Lemma 7.9. We need it later on due to a lack of zero boundary conditions.

Lemma 7.20. *Let $f \in L^2(Q)$ and $\theta \in (1/4, 1/2)$. Let $\widetilde{\Gamma}$ be the union of exact two of the sets Γ_1, Γ_2 and Γ_3, and $\widetilde{\Gamma}' = \Gamma \setminus \widetilde{\Gamma}$. Furthermore, let*

$$D := H^{3/2+\theta}(Q) \cap H^1_{\widetilde{\Gamma}}(Q)$$

and $g \in L^2(\widetilde{\Gamma}')$. Then there exists a unique function $v \in H^1_{\widetilde{\Gamma}}(Q)$ such that

$$\int_Q v\varphi \,\mathrm{d}x + \int_Q \nabla v \cdot \nabla\varphi \,\mathrm{d}x = \int_Q f\varphi \,\mathrm{d}x + \int_{\widetilde{\Gamma}'} g\varphi \,\mathrm{d}\sigma \tag{7.37}$$

for all $\varphi \in H^1_{\widetilde{\Gamma}}(Q)$. If $g \in H^\theta(\widetilde{\Gamma}')$, then we additionally have $v \in D$, $v - \Delta v = f$, $\partial_\nu v = g$ on $\widetilde{\Gamma}'$ and

$$\|v\|_{H^{3/2+\theta}} \le c\big(\|f\|_{L^2} + \|g\|_{H^\theta(\widetilde{\Gamma}')}\big).$$

Proof:
The proof works similar to the ones of Lemma 3.6 in [37] and of Lemma 7.9.

1) The Lemma of Lax-Milgram yields that problem (7.37) has a unique solution $\widetilde{u}$ $H^1_{\widetilde{\Gamma}}(Q)$. Let $L = \Delta$ be the Laplace operator on Q with Dirichlet boundary conditions on $\widetilde{\Gamma}$ and Neumann boundary conditions on $\widetilde{\Gamma}'$. It was shown in Lemma 3.6 of [37] that

$$D(L) = \{v \in H^2(Q) \cap H^1_{\widetilde{\Gamma}}(Q) \mid \partial_\nu v = 0 \text{ on } \widetilde{\Gamma}'\}.$$

Since L is m-accretive, we infer (after shifting L so that it is invertible) from Corollary 4.30 and 4.37 in [52] that

$$X_\alpha^L := D\big((I-L)^\alpha\big) = (L^2(Q), D(L))_{\alpha,2}$$

for all $\alpha \in (0,1)$. We interpolate for $\alpha \in (0,1) \setminus \{1/4, 3/4\}$ the inclusions

$$H_0^2(Q) \hookrightarrow D(L) \hookrightarrow H^2(Q)$$

and $L^2(Q) \to L^2(Q)$ to get

$$(L^2(Q), H_0^2(Q))_{\alpha,2} = H_0^{2\alpha}(Q) \hookrightarrow X_\alpha^L \hookrightarrow H^{2\alpha}(Q),$$

using Proposition 2.11 in [45]. Observe that $H_0^{2\alpha}(Q) = H^{2\alpha}(Q)$ for $\alpha \in (0,1/4)$, see Theorem 4.3.2.1 in [70]. This implies

$$X_\alpha^L = H^{2\alpha}(Q) = H_0^{2\alpha}(Q)$$

for $\alpha \in (0,1/4)$. Further, we have $X_{-\alpha}^L = \big(X_\alpha^L\big)^*$ due to the self-adjointness of L, see Proposition V.1.4.3 in [2]. This gives $X_\alpha^L = H^{-2\alpha}(Q)$ for $\alpha \in (-1/4, 0)$. The map $(I-L)^{-1} : X_{-\alpha}^L \to X_{1-\alpha}^L$ is continuous by Corollary V.1.3.9 in [2].

2) We assume without loss of generality that $\widetilde{\Gamma}' = \Gamma_1$. Let $R \subseteq \mathbb{R}^2$ be a rectangle that is congrent to one of the two congruent parts of Γ_1 and let Δ_R be the Dirichlet Laplacian on R with domain $D(\Delta_R) = H^2(R) \cap H_0^1(R)$. We conclude from Corollary 4.30 and 4.32 in [52] that

$$V_\alpha := D\big((-\Delta_R)^\alpha\big) = \big(L^2(R), D(-\Delta_R)\big)_{\alpha,2}.$$

Again from Corollary V.1.3.9 in [2] we infer $V_{-\alpha} = (V_\alpha)^*$. We see analogously to in part 1) that

$$H_0^{2\alpha}(R) \hookrightarrow V_\alpha \hookrightarrow H^{2\alpha}(R)$$

for $\alpha \in (0,1) \setminus \{1/4, 3/4\}$. This implies with Theorem 4.3.2.1 in [70] that

$$V_\alpha = H^{2\alpha}(R) = H_0^{2\alpha}(R)$$

for $\alpha \in (0,1/4)$. By duality, we also have $V_\alpha = H^{-2\alpha}(R)$ for $\alpha \in (-1/4, 0)$.

3) We denote by J_x and J_y the projections of R onto the x- and the y-axis, respectively, and set $H_D^{2\alpha} := H^{2\alpha} \cap H_0^{\min\{1,2\alpha\}}$ for all $\alpha \in (0,1]$. The operator Δ_R equals the sum of ∂_{xx} and ∂_{yy} with domains $H_D^2(J_x, L^2(J_y))$ and $H_D^2(J_y, L^2(J_x))$, respectively. As in the proof of Lemma 3.6 in [37] we see

$$D(\Delta_R) = D_0(\partial_{xx}) \cap D_0(\partial_{yy}) = H_D^2\big(J_x, L^2(J_y)\big) \cap H_D^2\big(J_y, L^2(J_x)\big),$$

where $D_0(\partial_{xx})$ and $D_0(\partial_{yy})$ are the domains of ∂_{xx} and ∂_{yy} with Dirichlet bonudary conditions, respectively. Due to [32] and the proof of Lemma 3.6 in [37] we hence have for $\alpha > 1/4$ that

$$V_\alpha = (L^2(Q), D_0(\partial_{xx}))_{\alpha,2} \cap (L^2(Q), D_0(\partial_{yy}))_{\alpha,2}$$

$$
\begin{aligned}
&= H^{2\alpha}_D(J_x, L^2(J_y)) \cap H^{2\alpha}_D(J_y, L^2(J_x)) \cap H^{2\alpha}(R) \\
&\subseteq \{u \in H^{2\alpha}(R) \mid \operatorname{tr} u = 0 \text{ on } \partial R\}.
\end{aligned}
$$

4) We assume $g \in C_c^\infty(\Gamma_1)$ and look at the two restrictions $g_1 \in C_c^\infty(\Gamma_1^-)$ and $g_2 \in C_c^\infty(\Gamma_1^+)$. We define w, $w^{(1)}$ and $w^{(2)}$ as in the proof of Lemma 7.9, i.e. for instance

$$
w^{(1)}(x_1, x_2, x_3) := -\big(\chi(x_1 - a_1^-)(-\Delta_R)^{-1/2} \exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big) g_1\big)(x_2, x_3)
$$

with a C^∞-function $\chi : [0, a_1^+ - a_1^-] \to \mathbb{R}$ with $\operatorname{supp}\chi \subseteq \big[0, \frac{1}{2}(a_1^+ - a_1^-)\big]$ and $\chi = 1$ on $\big[0, \frac{1}{4}(a_1^+ - a_1^-)\big]$. Further, we define for all $x_1 \in (a_1^-, a_1^+)$ the function $\psi(x_1) : R \to \mathbb{R}$ by

$$
\psi(x_1) := (-\Delta_R)^{-1/4} \exp\big((x_1 - a_1^-)(-\Delta_R)^{1/2}\big)(-\Delta_R)^{\theta/2} g_1
$$

and have the crucial estimate

$$
\|\psi(x_1)\|_{L^2(R)} \leq c(x_1) \|g_1\|_{H^\theta(R)},
$$

which holds true due to Proposition 6.2 in [52]. Taking the derivatives of $w^{(1)}$ and rearranging the operators gives

$$
\begin{aligned}
w^{(1)}(x_1, x_2, x_3) &= -\big(\chi(x_1 - a_1^-)(-\Delta_R)^{-3/4-\theta/2}\psi(x_1)\big)(x_2, x_3), \\
\partial_1 w^{(1)}(x_1, x_2, x_3) &= -\big(\chi(x_1 - a_1^-)(-\Delta_R)^{-1/4-\theta/2}\psi(x_1)\big)(x_2, x_3) \\
&\quad - \big(\chi'(x_1 - a_1^-)(-\Delta_R)^{-3/4-\theta/2}\psi(x_1)\big)(x_2, x_3), \\
\partial_k w^{(1)}(x_1, x_2, x_3) &= -\big(\chi(x_1 - a_1^-)\partial_k(-\Delta_R)^{-3/4/-\theta/2}\psi(x_1)\big)(x_2, x_3), \\
\partial_{11} w^{(1)}(x_1, x_2, x_3) &= -\big(\chi(x_1 - a_1^-)(-\Delta_R)^{1/4-\theta/2}\psi(x_1)\big)(x_2, x_3) \\
&\quad - 2\big(\chi'(x_1 - a_1^-)(-\Delta_R)^{-1/4-\theta/2}\psi(x_1)\big)(x_2, x_3) \\
&\quad - \big(\chi''(x_1 - a_1^-)(-\Delta_R)^{-3/4-\theta/2}\psi(x_1)\big)(x_2, x_3), \\
\partial_{1k} w^{(1)}(x_1, x_2, x_3) &= -\big(\chi(x_1 - a_1^-)\partial_k(-\Delta_R)^{-1/4-\theta/2}\psi(x_1)\big)(x_2, x_3) \\
&\quad - \big(\chi'(x_1 - a_1^-)\partial_k(-\Delta_R)^{-3/4-\theta/2}\psi(x_1)\big)(x_2, x_3), \\
\partial_{kl} w^{(1)}(x_1, x_2, x_3) &= -\big(\chi(x_1 - a_1^-)\partial_{kl}(-\Delta_R)^{-3/4-\theta/2}\psi(x_1)\big)(x_2, x_3)
\end{aligned}
$$

for all $k, l \in \{2, 3\}$. Due to $1/4 - \theta/2 \in (-1/4, 0)$, the operator $(-\Delta_R)^{1/4-\theta/2}_{-1}$ is continuous from $L^2(R)$ to $V_{\theta/2-1/4}$ and we have $V_{\theta/2-1/4} = H^{\theta-1/2}(R)$. We furthermore have

$$
\begin{aligned}
\partial_k V_{1/4+\theta/2} &\subseteq \partial_k H^{1/2+\theta}(R) \subseteq H^{\theta-1/2}(R), \\
\partial_{kl} V_{3/4+\theta/2} &\subseteq \partial_{kl} H^{3/2+\theta}(R) \subseteq H^{\theta-1/2}(R)
\end{aligned}
$$

for all $k, l \in \{2, 3\}$. The other appearing terms are of the same type or even regular. We thus infer with the boundedness of χ and its derivative that $w^{(1)}$ belongs to $H^{3/2+\theta}(Q)$. Arguing in the same way with $w^{(2)}$ we obtain that w belongs to $H^{3/2+\theta}(Q)$.

Set $\widetilde{f} := f - w + \Delta w$. Since $\theta/2 - 1/4 \in (-1/4, 0)$, we conclude that $\widetilde{f}$ is contained in $H^{\theta-1/2}(Q) = X^L_{\theta/2-1/4}$. Together with $w \in H^{3/2+\theta}(Q)$ we thus infer that

$$u := (I - L)^{-1}\widetilde{f} + w$$

belongs to $X^L_{3/4+\theta/2} \subseteq H^{3/2+\theta}(Q)$. Furthermore, $u - \Delta u = f$, $u = 0$ on $\widetilde{\Gamma}$ and $\partial_\nu u = g$ on $\widetilde{\Gamma}'$.

5) We now approximate $g \in H^\theta(\Gamma_1)$ by a sequence $(g_n)_{n\in\mathbb{N}}$ in $C_c^\infty(\Gamma_1)$ with $g_n \to g$ in $H^\theta(\Gamma_1)$ as $n \to \infty$, which is possible due to Proposition 1.17 in [52]. We take a corresponding sequence $(u_n)_{n\in\mathbb{N}}$ in $H^{3/2+\theta}(Q)$ from step 3) and obtain with the same estimates as above that

$$\|u_n - u_m\|_{H^{3/2+\theta}} \leq c\,\|g_n - g_m\|_{H^\theta(\Gamma_1)} \longrightarrow 0$$

as $n, m \to \infty$. Hence, (u_n) has a limit u in $H^{3/2+\theta}(Q)$. The continuity of the Dirichlet trace map and the Neumann trace map implies $u = 0$ on $\Gamma \setminus \Gamma_1$ and $\partial_\nu u = g$ on Γ_1.

By the divergence theorem one checks that u satisfies (7.37) for all $\varphi \in H^1_{\widetilde{\Gamma}}(Q)$, so that it is equal to $\widetilde{u}$ from step 1). □

We now state an analogon of the above result in the spaces X_0, $X^{(0)}_{\mathrm{div}}$ and $X^{(2)}_{\mathrm{div}}$.

Proposition 7.21. *(a) Let $\sigma = 0$. Then the operator M_0 generates a unitary C_0-semigroup e^{tM_0} on X_0. For $(\boldsymbol{E}_0, \boldsymbol{H}_0) \in D(M_0)$ and $\boldsymbol{J}_0 = 0$, the function*

$$(\boldsymbol{E}(t), \boldsymbol{H}(t)) := e^{tM_0}(\boldsymbol{E}_0, \boldsymbol{H}_0)$$

is the unique solution to (7.1) in $C^1([0,\infty), X_0) \cap C([0,\infty), D(M_0))$, where $\rho = 0$.

(b) The operator $M^{(0)}_{\mathrm{div}}$ generates a C_0-semigroup $e^{tM^{(0)}_{\mathrm{div}}}$ on $X^{(0)}_{\mathrm{div}}$. For $(\boldsymbol{E}_0, \boldsymbol{H}_0) \in D(M^{(0)}_{\mathrm{div}})$ and $(\boldsymbol{J}_0, 0) \in C\big([0,\infty), D(M^{(0)}_{\mathrm{div}})\big) + C^1\big([0,\infty), X^{(0)}_{\mathrm{div}}\big)$, the function

$$(\boldsymbol{E}(t), \boldsymbol{H}(t)) := e^{tM^{(0)}_{\mathrm{div}}}(\boldsymbol{E}_0, \boldsymbol{H}_0) - \int_0^t e^{(t-s)M^{(0)}_{\mathrm{div}}}\big(\tfrac{1}{\varepsilon}\boldsymbol{J}_0(s), 0\big)\,\mathrm{d}s, \quad t \geq 0$$

is the unique solution to (7.1) in $C^1\big([0,\infty), X^{(0)}_{\mathrm{div}}\big) \cap C\big([0,\infty), D(M^{(0)}_{\mathrm{div}})\big)$ with

$$\begin{aligned}\rho = \operatorname{div}(\varepsilon\boldsymbol{E}(t)) &= \operatorname{div}(\varepsilon\boldsymbol{E}_0) - \int_0^t \operatorname{div}(\sigma\boldsymbol{E}(s) + \boldsymbol{J}_0(s))\,\mathrm{d}s \\ &= e^{-\frac{\sigma}{\varepsilon}t}\operatorname{div}(\varepsilon\boldsymbol{E}_0) - \int_0^t e^{-\frac{\sigma}{\varepsilon}(t-s)}\Big(\nabla\big(\tfrac{\sigma}{\varepsilon}\big)\cdot\varepsilon\boldsymbol{E}(s) + \operatorname{div}\boldsymbol{J}_0(s)\Big)\,\mathrm{d}s\end{aligned} \tag{7.38}$$

in $L^2(Q)$ for $t \geq 0$. The semigroup satisfies

$$\Big\|e^{tM^{(0)}_{\mathrm{div}}}(\boldsymbol{E}_0, \boldsymbol{H}_0)\Big\|_{X^{(0)}_{\mathrm{div}}} \leq c(1+t)\,\|(\boldsymbol{E}_0, \boldsymbol{H}_0)\|_{X^{(0)}_{\mathrm{div}}}$$

for all $t \geq 0$ and $(\boldsymbol{E}_0, \boldsymbol{H}_0) \in X^{(0)}_{\mathrm{div}}$.

(c) Let $\varepsilon, \mu, \sigma \in W^{2,3}(Q)$. Then $M_{\mathrm{div}}^{(2)}$ generates a C_0-semigroup $e^{tM_{\mathrm{div}}^{(2)}}$ on $X_{\mathrm{div}}^{(2)}$. For $(\boldsymbol{E}_0, \boldsymbol{H}_0) \in D(M_{\mathrm{div}}^{(2)})$ and $\big(\frac{1}{\varepsilon}\boldsymbol{J}_0, 0\big) \in C\big([0,\infty), D(M_{\mathrm{div}}^{(2)})\big) + C^1\big([0,\infty), X_{\mathrm{div}}^{(2)}\big)$, the function

$$(\boldsymbol{E}(t), \boldsymbol{H}(t)) := e^{tM_{\mathrm{div}}^{(2)}}(\boldsymbol{E}_0, \boldsymbol{H}_0) - \int_0^t e^{(t-s)M_{\mathrm{div}}^{(2)}} \big(\tfrac{1}{\varepsilon}\boldsymbol{J}_0(s), 0\big)\, \mathrm{d}s, \quad t \geq 0$$

is the unique solution to (7.1) in $C^1\big([0,\infty), X_{\mathrm{div}}^{(2)}\big) \cap C\big([0,\infty), D(M_{\mathrm{div}}^{(2)})\big)$, where ρ is given as in (7.38). The semigroup satisfies

$$\Big\| e^{tM_{\mathrm{div}}^{(2)}}(\boldsymbol{E}_0, \boldsymbol{H}_0) \Big\|_{X_{\mathrm{div}}^{(2)}} \leq c\big(1+t^2\big)\, \|(\boldsymbol{E}_0, \boldsymbol{H}_0)\|_{X_{\mathrm{div}}^{(2)}}$$

for all $t \geq 0$ and $(\boldsymbol{E}_0, \boldsymbol{H}_0) \in X_{\mathrm{div}}^{(2)}$.

All three semigroups are restrictions of e^{tM}.

Proof:

Part (a) was shown in Proposition 3.5 of [37].

(b) Let $(\mathbf{E}_0, \mathbf{H}_0) \in X_{\mathrm{div}}^{(0)}$ and $t \geq 0$. Set

$$(\widetilde{\mathbf{E}}(t), \widetilde{\mathbf{H}}(t)) := e^{tM}(\mathbf{E}_0, \mathbf{H}_0).$$

Proposition 7.19 shows that $\Big\|(\widetilde{\mathbf{E}}(t), \widetilde{\mathbf{H}}(t))\Big\|_X \leq \|(\mathbf{E}_0, \mathbf{H}_0)\|_X$. Hence, formula (7.36b) yields that $\operatorname{div}(\varepsilon\widetilde{\mathbf{E}}(t))$ belongs to $L^2(Q)$ and that

$$\Big\|\operatorname{div}(\varepsilon\widetilde{\mathbf{E}}(t))\Big\|_{L^2(Q)} \leq \|\operatorname{div}(\varepsilon\mathbf{E}_0)\|_{L^2(Q)} + ct\, \|(\mathbf{E}_0, \mathbf{H}_0)\|_X\,.$$

Moreover, $\operatorname{div}(\varepsilon\mathbf{E}(t))$ tends to $\operatorname{div}(\varepsilon\mathbf{E}_0)$ in $L^2(Q)$ as $t \to 0$. The magnetic conditions in $X_{\mathrm{div}}^{(0)}$ are satisfied by $\mathbf{H}(t)$ due to (7.36d) and (7.36e). The semigroup e^{tM} thus leaves $X_{\mathrm{div}}^{(0)}$ invariant and is strongly continuous on this space. Hence, it satisfies the asserted estimate due to Section II.2.3 in [23]. Here we use that $M_{\mathrm{div}}^{(0)}$ is the part of M in $X_{\mathrm{div}}^{(0)}$ by Lemma 7.12 and $\big(\frac{1}{\varepsilon}\mathbf{J}_0, 0\big)$ belongs to $C\big([0,\infty), D(M_{\mathrm{div}}^{(0)})\big) + C^1\big([0,\infty), X_{\mathrm{div}}^{(0)}\big)$.

(c) 1) We now take $(\mathbf{E}_0, \mathbf{H}_0) \in X_{\mathrm{div}}^{(2)}$ and define $(\widetilde{\mathbf{E}}(t), \widetilde{\mathbf{H}}(t))$ as in the proof of part (b). The strategy of the proof is again to check that the semigroup e^{tM} leaves $X_{\mathrm{div}}^{(2)}$ invariant and is strongly continuous thereon. Then we conclude the assertion by Section II.2.3 in [23].

2) As above the magnetic field $\widetilde{\mathbf{H}}(t)$ satisfies the divergence and boundary conditions in $X_{\mathrm{div}}^{(2)}$ for $t \geq 0$ and the map $t \mapsto \operatorname{div}(\varepsilon\widetilde{\mathbf{E}}(t))$ is continuous in $L^2(Q)$. Moreover, $t \mapsto (\widetilde{\mathbf{E}}(t), \widetilde{\mathbf{H}}(t))$ is continuous in the space $D(M^2)$, so that $\widetilde{\mathbf{E}}$ is contained in $C([0,\infty), H_0(\operatorname{curl}, Q))$. Taking also the identity

$$\operatorname{div}(\widetilde{\mathbf{E}}(t)) = \operatorname{div}\big(\tfrac{1}{\varepsilon}\varepsilon\widetilde{\mathbf{E}}(t)\big) = \nabla\big(\tfrac{1}{\varepsilon}\big) \cdot \varepsilon\widetilde{\mathbf{E}}(t) + \frac{1}{\varepsilon}\operatorname{div}(\varepsilon\widetilde{\mathbf{E}}(t))$$

for $t \geq 0$ into account, we see that $\widetilde{\mathbf{E}}$ belongs to $C([0,\infty), H(\mathrm{div}, Q))$. Proposition 7.4 thus shows that $\widetilde{\mathbf{E}}$ is a continuous map into $H^1(Q)^3$ and

$$\left\|\widetilde{\mathbf{E}}(t)\right\|_{H^1} \leq c(1+t)\,\|(\mathbf{E}_0, \mathbf{H}_0)\|_{X^{(2)}_{\mathrm{div}}}\,.$$

3) This fact allows us to differentiate (7.36b) in $L^2(Q)$, obtaining

$$\begin{aligned}\nabla \,\mathrm{div}(\varepsilon\widetilde{\mathbf{E}}(t)) = {} & -te^{-\frac{\sigma}{\varepsilon}t}\nabla\big(\tfrac{\sigma}{\varepsilon}\big)\,\mathrm{div}(\varepsilon\mathbf{E}_0) + e^{-\frac{\sigma}{\varepsilon}t}\nabla\,\mathrm{div}(\varepsilon\mathbf{E}_0) \\ & + \int_0^t e^{-\frac{\sigma}{\varepsilon}(t-s)}\Big((t-s)\big(\nabla\big(\tfrac{\sigma}{\varepsilon}\big)\big)^2\varepsilon\widetilde{\mathbf{E}}(s) + D^2\big(\tfrac{\sigma}{\varepsilon}\big)\varepsilon\widetilde{\mathbf{E}}(s) \\ & \quad - \nabla\big(\tfrac{\sigma}{\varepsilon}\big)(\nabla\varepsilon)\widetilde{\mathbf{E}}(s) - \varepsilon\big(\partial\widetilde{\mathbf{E}}(s)\big)^T\nabla\big(\tfrac{\sigma}{\varepsilon}\big)\Big)\,\mathrm{d}s\end{aligned}$$

for all $t \geq 0$, where D^2u denotes the matrix with the second derivatives of a function u. Using the properties of σ and ε, the H^1-continuity of $\widetilde{\mathbf{E}}$ and the embedding $H^1(Q)^3 \hookrightarrow L^6(Q)^3$, we conclude that $\mathrm{div}(\varepsilon\widetilde{\mathbf{E}}(t))$ belongs to $C([0,\infty), L^2(Q))$ and fulfils the estimate

$$\begin{aligned}\left\|\nabla\,\mathrm{div}(\varepsilon\widetilde{\mathbf{E}}(t))\right\|_{L^2} = {} & t\left\|\nabla\big(\tfrac{\sigma}{\varepsilon}\big)\right\|_{L^\infty}\|\mathrm{div}(\varepsilon\mathbf{E}_0)\|_{L^2} + \|\nabla\,\mathrm{div}(\varepsilon\mathbf{E}_0)\|_{L^2} \\ & + \int_0^t\Big((t-s)\left\|\nabla\big(\tfrac{\sigma}{\varepsilon}\big)\right\|_{L^\infty}^2\|\varepsilon\|_{L^\infty}\left\|\widetilde{\mathbf{E}}(s)\right\|_{L^2} \\ & \quad + \left\|D^2\big(\tfrac{\sigma}{\varepsilon}\big)\right\|_{L^3}\|\varepsilon\|_{L^\infty}\left\|\widetilde{\mathbf{E}}(s)\right\|_{H^1} + \left\|\nabla\big(\tfrac{\sigma}{\varepsilon}\big)\right\|_{L^\infty}\|\nabla\varepsilon\|_{L^\infty}\left\|\widetilde{\mathbf{E}}(s)\right\|_{L^2} \\ & \quad + \left\|\nabla\big(\tfrac{\sigma}{\varepsilon}\big)\right\|_{L^\infty}\|\varepsilon\|_{L^\infty}\left\|\nabla\widetilde{\mathbf{E}}(s)\right\|_{L^2}\Big)\,\mathrm{d}s \\ \leq {} & c\Big(t\,\|\mathrm{div}(\varepsilon\mathbf{E}_0)\|_{H^1} + (t+t^2)\left\|\widetilde{\mathbf{E}}(t)\right\|_{L^2} + t\left\|\widetilde{\mathbf{E}}(t)\right\|_{H^1}\Big).\end{aligned}$$

Together with step 1), we thus deduce the continuity of $t \mapsto \mathrm{div}(\varepsilon\widetilde{\mathbf{E}}(t))$ in $H^1(Q)$ and the bound

$$\left\|\mathrm{div}(\varepsilon\widetilde{\mathbf{E}}(t))\right\|_{H^1(Q)} \leq c(1+t^2)\,\|(\mathbf{E}_0, \mathbf{H}_0)\|_{X^{(2)}_{\mathrm{div}}}\,.$$

4) We still have to show the continuity of $t \mapsto \mathrm{div}(\varepsilon\widetilde{\mathbf{E}}(t))$ in the smaller space $H^1_{00}(Q)$. To this aim we first show that $t \mapsto \widetilde{\mathbf{E}}(t)$ is continuous with values in $H^{15/8}(Q)$, which will allow us to take traces on the edges of Q. Let $t \geq 0$. As in the proof of Proposition 7.11 we get

$$\begin{aligned}\|\Delta\mathbf{E}(t)\|_{L^2}\,, \|\Delta\mathbf{H}(t)\|_{L^2} &\leq c\big(\left\|M^2(\mathbf{E}(t), \mathbf{H}(t))\right\|_X + \|(\mathbf{E}(t), \mathbf{H}(t))\|_{X^{(0)}_{\mathrm{div}}}\big) \\ &\leq c\,\|(\mathbf{E}(t), \mathbf{H}(t))\|_{X^{(2)}_{\mathrm{div}}}\,,\end{aligned}$$

as well as $\Delta(\varepsilon E_1(t)) \in L^2(Q)$ and $\varepsilon E_1(t) \in H^2_{loc}(Q)$. As therein we get equation (7.19) with $\rho(t) \in H^{1/2}(\Gamma) \hookrightarrow H^\theta(\Gamma)$ for $\theta = 3/8 \in (1/4, 1/2)$. Lemma 7.20 hence yields $\varepsilon E_1(t) \in H^{15/8}(Q)$. Due to the assumptions on ε, the multiplication operator $f \mapsto \frac{1}{\varepsilon}f$ is continuous from $H^1(Q)$ to $H^1(Q)$ and from $H^2(Q)$ to $H^2(Q)$. By interpolation it is

thus also continuous from $H^{15/8}(Q)$ to $H^{15/8}(Q)$, from which we infer $E_1(t) \in H^{15/8}(Q)$. This yields $\operatorname{tr}(E_1(t)) \in H^\alpha(\Gamma)$ for all $\alpha \in (1/2, 1)$ by Theorem 3.1 in [45]. The trace $\operatorname{tr}(E_1(t)) = 0$ on $\Gamma_2 \cup \Gamma_3$ now gives by the same theorem that $\operatorname{tr}(E_1(t)) = 0$ on $\partial\widehat{\Gamma}$ for all faces $\widehat{\Gamma} \subseteq \Gamma$ (in $H^{\alpha-1/2}(\widehat{\Gamma})$ for all $\alpha \in (1/2, 1)$). So, $\operatorname{tr}(\varepsilon E_1(t)) = 0$ on all faces $\widehat{\Gamma} \subseteq \Gamma$. Analogous results hold for $E_2(t)$ and $E_3(t)$.

5) The function $e^{-\frac{\sigma}{\varepsilon}t}$ is continuous from $L^2(\widehat{\Gamma})$ to $L^2(\widehat{\Gamma})$ and from $H_0^1(\widehat{\Gamma})$ to $H_0^1(\widehat{\Gamma})$ for all faces $\widehat{\Gamma}$ of Q. So, it is also continuous from $H_0^{1/2}(\widehat{\Gamma})$ to $H_0^{1/2}(\widehat{\Gamma})$, which yields that $e^{-\frac{\sigma}{\varepsilon}t}\operatorname{div}(\varepsilon\mathbf{E}_0)$ is contained in $H_0^{1/2}(\widehat{\Gamma})$ for all faces $\widehat{\Gamma}$ of Q. Thus, $t \mapsto e^{-\frac{\sigma}{\varepsilon}t}\operatorname{div}(\varepsilon\mathbf{E}_0)$ with values in $H_{00}^1(Q)$. We have $\nabla\left(\frac{\sigma}{\varepsilon}\right) \in W^{1,3}(Q)$. Let φ_n be functions in $C^\infty(\overline{Q})$ with $\varphi_n \to \nabla\left(\frac{\sigma}{\varepsilon}\right)$ in $W^{1,3}(Q)$. Then

$$\operatorname{tr}(\varphi_n \varepsilon \mathbf{E}(s)) = \operatorname{tr}(\varphi_n)\operatorname{tr}(\varepsilon\mathbf{E}(s)) = 0$$

on $\partial\widehat{\Gamma}$ for all faces $\widehat{\Gamma}$ of Q. Taking the limit we get that $\nabla\left(\frac{\sigma}{\varepsilon}\right) \cdot \varepsilon\mathbf{E}(s)$ vanishes on all edges of Q. We apply the Sobolev embedding $H^{15/8}(Q) \hookrightarrow L^\infty(Q)$ from Theorem 4.6.1 in [70] to $\mathbf{E}(s)$ and get with $\varepsilon \in L^\infty(Q)$ and $\nabla\left(\frac{\sigma}{\varepsilon}\right) \in W^{1,3}(Q)$ that

$$\int_0^t \nabla\left(\tfrac{\sigma}{\varepsilon}\right) \cdot \varepsilon\mathbf{E}(s)\,\mathrm{d}s$$

belongs to $W^{1,3}(Q)$, so that its trace belongs to $W_0^{2/3,3}(\Gamma)$ due to the vanishing trace of the integrand. From the inclusions $L^3(\Gamma) \hookrightarrow L^2(\Gamma)$ and $W_0^{1,3}(\Gamma) \hookrightarrow H_0^1(\Gamma)$ we infer by interpolating the embedding $W_0^{2/3,3}(\Gamma) \hookrightarrow H_0^{1/2}(\Gamma)$. Thus, we get altogether with (7.36b) that ρ is continuous with values in $H_0^{1/2}(\Gamma)$ and hence that ρ is continuous with values in $H_{00}^1(Q)$. □

8. The ADI splitting scheme and properties of the splitting operators

This chapter is devoted to the splitting operators and the splitting scheme we construct with them. We show in the Sections 8.1, 8.2 and 8.3 that the splitting operators in the L^2-setting and their restrictions to the subspace of H^1 and H^2 generate quasicontractive strongly continuous semigroups, respectively. This implies crucial estimate of their resolvents. After presenting the ADI splitting scheme in Section 8.4, we explain its efficiency in Section 8.5.

8.1. Properties of the splitting operators in the L^2-setting

We start with a basic result in X.

Proposition 8.1. *(a) The operators A and B generate C_0-semigroups of contractions on X. In particular,*

$$\left\|(I-\tau A)^{-1}\right\|_{\mathcal{B}(X)} \leq 1 \qquad \textit{and} \qquad \left\|(I-\tau B)^{-1}\right\|_{\mathcal{B}(X)} \leq 1$$

for all $\tau > 0$.

(b) For all $\tau > 0$ we have

$$\left\|(I+\tau A)(I-\tau A)^{-1}\right\|_{\mathcal{B}(X)} \leq 1 \qquad \textit{and} \qquad \left\|(I+\tau B)(I-\tau B)^{-1}\right\|_{\mathcal{B}(X)} \leq 1.$$

Proof:
(a) In Lemma 4.3 in [37] it was shown that A_0 and B_0, being defined in (7.24), are skew-adjoint on X. Therefore, they generate by Stone's Theorem C_0-semigroups (even C_0-groups) of unitary operators on X. Due to $-\sigma \leq 0$ and the boundedness of σ, Theorem III.2.7 in [23] shows that the operators A and B, see (7.23), also generate C_0-semigroups of contractions on X. In particular, all $\lambda > 0$ are in the resolvent set of A

and in the resolvent set of B. For $\tau > 0$ the Theorem of Hille–Yosida gives

$$\left\|(I-\tau A)^{-1}x\right\|_X = \frac{1}{\tau}\left\|\left(\tfrac{1}{\tau}I - A\right)^{-1}x\right\|_X \leq \frac{1}{\tau}\cdot\tau\,\|x\|_X = \|x\|_X$$

for all $x \in X$, which yields the desired resolvent estimate. The operator B is treated in the same way.
(b) Let $\tau > 0$. For all $x \in D(A)$ we have by the dissipativity of A, see the remark to Assumption 7 in [59],

$$\begin{aligned}\|(I+\tau A)x\|_X^2 &= \|x\|_X^2 + 2\,\mathrm{Re}(\tau Ax \mid x)_X + \|\tau Ax\|_X^2 \\ &\leq \|x\|_X^2 - 2\,\mathrm{Re}(\tau Ax \mid x)_X + \|\tau Ax\|_X^2 = \|(I-\tau A)x\|_X^2\,.\end{aligned}$$

Because each $x \in D(A)$ can be written as $x = (I-\tau A)^{-1}y$ for some $y \in X$, we thus have

$$\left\|(I+\tau A)(I-\tau A)^{-1}y\right\|_X \leq \|y\|_X\,.$$

Hence, with the same argumentation for B, we infer

$$\left\|(I+\tau A)(I-\tau A)^{-1}\right\|_{\mathcal{B}(X)} \leq 1 \quad \text{and} \quad \left\|(I+\tau B)(I-\tau B)^{-1}\right\|_{\mathcal{B}(X)} \leq 1,$$

which finishes the proof. □

8.2. Properties of the splitting operators in the H^1-setting

We conclude the following corollary on first-order traces by Lemma 7.10 from the zero-order boundary conditions of $D(A_Y)$ and $D(B_Y)$, see Subsection 7.2.2 for the definition of the operators A_Y and B_Y.

Corollary 8.2. *(a) Let $(u,v) \in D(A_Y)$. Then*

$$\begin{aligned}\partial_3 u_2 = \partial_2 u_3 = \partial_3 u_3 = \partial_3 v_1 = 0 \qquad &\text{on } \Gamma_1,\\ \partial_1 u_3 = \partial_3 u_1 = \partial_1 u_1 = \partial_1 v_2 = 0 \qquad &\text{on } \Gamma_2,\\ \partial_2 u_1 = \partial_1 u_2 = \partial_2 u_2 = \partial_2 v_3 = 0 \qquad &\text{on } \Gamma_3.\end{aligned}$$

(b) Let $(u,v) \in D(B_Y)$. Then

$$\begin{aligned}\partial_2 u_3 = \partial_2 u_2 = \partial_3 u_2 = \partial_2 v_1 = 0 \qquad &\text{on } \Gamma_1,\\ \partial_3 u_1 = \partial_3 u_3 = \partial_1 u_3 = \partial_3 v_2 = 0 \qquad &\text{on } \Gamma_2,\\ \partial_1 u_2 = \partial_1 u_1 = \partial_2 u_1 = \partial_1 v_3 = 0 \qquad &\text{on } \Gamma_3.\end{aligned}$$

The above first-order boundary conditions can be used to rewrite the domains of A_Y and B_Y.

Corollary 8.3. *We have*

$$\begin{aligned} D(A_Y) = \{(u,v) \in H^1(Q)^6 \mid\ & u_j = 0 \text{ on } \Gamma \setminus \Gamma_j,\ \ v_j = 0 \text{ on } \Gamma_j, \text{ for all } j \in \{1,2,3\}, \\ & \partial_2 u_1, \partial_3 u_2, \partial_1 u_3, \partial_3 v_1, \partial_1 v_2, \partial_2 v_3 \in H^1(Q), \\ & \partial_3 v_1 = 0 \text{ on } \Gamma_3,\ \ \partial_1 v_2 = 0 \text{ on } \Gamma_1,\ \ \partial_2 v_3 = 0 \text{ on } \Gamma_2\} \end{aligned}$$

and

$$\begin{aligned} D(B_Y) = \{(u,v) \in H^1(Q)^6 \mid\ & u_j = 0 \text{ on } \Gamma \setminus \Gamma_j,\ \ v_j = 0 \text{ on } \Gamma_j, \text{ for all } j \in \{1,2,3\}, \\ & \partial_3 u_1, \partial_1 u_2, \partial_2 u_3, \partial_2 v_1, \partial_3 v_2, \partial_1 v_3 \in H^1(Q), \\ & \partial_2 v_1 = 0 \text{ on } \Gamma_2,\ \ \partial_3 v_2 = 0 \text{ on } \Gamma_3,\ \ \partial_1 v_3 = 0 \text{ on } \Gamma_1\}. \end{aligned}$$

In the next lemmas we collect some basic properties of A_Y and B_Y.

Lemma 8.4. *The operators A_Y and B_Y are closed in Y and densely defined on Y.*

Proof:
1) To show the closedness of A_Y we take a sequence $(u_n, v_n)_{n\in\mathbb{N}} \subseteq D(A_Y)$ with $(u_n, v_n) \to (u, v)$ in Y and $A(u_n, v_n) \to (f, g)$ in Y as $n \to \infty$. Then (u, v) fulfils the zero-order boundary conditions of Y by the continuity of the occurring traces. Moreover, $C_2 u_n$ and $C_1 v_n$ tend to $C_2 u$ and $C_1 v$ in $L^2(Q)^3$, respectively, as $n \to \infty$, and $A(u, v) = (f, g)$. Additionally, we deduce from Lemma 7.2 that $\frac{\sigma}{2} u_n \to \frac{\sigma}{2} u$ in $H^1(Q)^3$ and

$$C_1 v_n = \left(-\tfrac{\sigma}{2} u_n + C_1 v_n\right) + \tfrac{\sigma}{2} u_n \longrightarrow \varepsilon f + \tfrac{\sigma}{2} u = C_1 v$$

and $C_2 u_n \to \mu g$ in $H^1(Q)^3$ as $n \to \infty$. As a result, $C_2 u$ and $C_1 v$ belong to $H^1(Q)^3$ and (u, v) satisfies the first-order boundary conditions of $D(A_Y)$. Altogether we have $(u, v) \in D(A_Y)$ and $A(u, v) = (f, g)$.

To show the closedness of B_Y we take a sequence $(u_n, v_n)_{n\in\mathbb{N}} \subseteq D(B_Y)$ with $(u_n, v_n) \to (u, v)$ in Y and $B(u_n, v_n) \to (f, g)$ in Y as $n \to \infty$. Then (u, v) fulfils the zero-order boundary conditions of Y. Additionally, $C_1 u_n$ and $C_2 v_n$ converge to $C_1 u$ and $C_2 v$ in $H^1(Q)^3$, respectively, as $n \to \infty$, and $B(u, v) = (f, g)$. Furthermore, we have $\frac{\sigma}{2} u_n \to \frac{\sigma}{2} u$ in $H^1(Q)^3$ and

$$-C_2 v_n = \left(-\tfrac{\sigma}{2} u_n - C_2 v_n\right) + \tfrac{\sigma}{2} u_n \longrightarrow \varepsilon f + \tfrac{\sigma}{2} u$$

and $-C_1 u_n \to \mu g$ in $H^1(Q)^3$ as $n \to \infty$. This yields that $C_1 u$ and $C_2 v$ belong to $H^1(Q)^3$ and that (u, v) fulfils the first-order boundary conditions of $D(B_Y)$. Hence, we have $(u, v) \in D(B_Y)$ and $B(u, v) = (f, g)$.

2) Let $(u, v) \in Y$, choose $n_0 \in \mathbb{N}$ with $n_0 \geq \frac{4}{d_{\min}}$ and let $n \geq n_0$. Let $\chi_n^{(j)}$ be the cut-off functions from the proof of Lemma 7.10, see (7.13), acting on the j-th variable. In the

sequel we use the standard C^∞-mollifiers $\rho_n^{(j)}$ with support in $[-\frac{1}{2n}, \frac{1}{2n}]$, acting on the j-th variable. We extend u_1 by 0 outside of Q and define the convolution

$$u_{1,n} := \big(\rho_n^{(2)} * \big(\chi_n^{(2)}\chi_n^{(3)}u_1\big)\big)|_Q.$$

The support of this function has a distance of at least $\frac{1}{2n}$ to $\Gamma_2 \cup \Gamma_3$, which implies the boundary condition on the first of the six components of elements of $D(A_Y)$, see Corollary 8.3. It is clear that $u_{1,n}$ and $\partial_1 u_{1,n}$ belong to $L^2(Q)$ and, letting the derivative act on the mollifier, also $\partial_2 u_{1,n}$. Due to $\chi_n^{(2)}, \chi_n^{(3)}, (\chi_n^{(3)})' \in L^\infty(Q)$, we have

$$\partial_3 u_{1,n} = \rho_n^{(2)} * \Big(\chi_n^{(2)}(\chi_n^{(3)})' u_1 + \chi_n^{(2)}\chi_n^{(3)}\partial_3 u_1\Big) \in L^2(Q),$$

so that together $u_{1,n} \in H^1(Q)$. From

$$\partial_{22} u_{1,n} = \partial_2\rho_n^{(2)} * \Big((\chi_n^{(2)})'\chi_n^{(3)}u_1 + \chi_n^{(2)}\chi_n^{(3)}\partial_2 u_1\Big)$$

and

$$\partial_{j2} u_{1,n} = \partial_2\rho_n^{(2)} * \partial_j\big(\chi_n^{(2)}\chi_n^{(3)}u_1\big)$$

for $j \in \{1,3\}$ we deduce with the same arguments that $\partial_2 u_{1,n}$ is contained in $H^1(Q)$. Standard results on mollifiers yield that

$$u_{1,n} \longrightarrow u_1 \quad \text{in } L^2(Q) \qquad \text{and} \qquad \partial_1 u_{1,n} \longrightarrow \partial_1 u_1 \quad \text{in } L^2(Q)$$

as $n \to \infty$. We argue analogously to the procedure in the proof of Lemma 7.10 (and with the notation from there). We conclude $u_1(x) = \int_{a_2^-}^{x_2} \partial_2 u_1(x_1, t, x_3)\,\mathrm{d}t$ from $u_1 = 0$ on Γ_2 and thus

$$\|u_1(\cdot, x_2, \cdot)\|_{L^2(Q_2)} \le (x_2 - a_2^-)^{1/2}\Big(\int_{a_2^-}^{x_2}\int_{Q_2} |\partial_2 u_1(x_2, t, x_3)|^2 \,\mathrm{d}(x_1, x_3)\,\mathrm{d}t\Big)^{1/2},$$

as well as

$$\|u_1(\cdot, x_2, \cdot)\|_{L^2(Q_2)} \le (a_2^+ - x_2)^{1/2}\Big(\int_{x_2}^{a_2^+}\int_{Q_2} |\partial_2 u_1(x_1, t, x_3)|^2 \,\mathrm{d}(x_1, x_3)\,\mathrm{d}t\Big)^{1/2}$$

for almost all $x_2 \in (a_2^-, a_2^+)$, so that

$$\begin{aligned}
\big\|(\chi_n^{(2)})'\chi_n^{(3)}u_1\big\|_{L^2} &\le n\Big(\int_{[a_2^-, a_2^- + \frac{2}{n}] \cup [a_2^+ - \frac{2}{n}, a_2^+]} \|u_1(\cdot, x_2, \cdot)\|_{L^2(Q_2)}^2 \,\mathrm{d}x_2\Big)^{1/2} \\
&\le 2\sqrt{n} \sup_{x_2 \in [a_2^-, a_2^- + \frac{2}{n}] \cup [a_2^+ - \frac{2}{n}, a_2^+]} \|u_1(\cdot, x_2, \cdot)\|_{L^2(Q_2)} \qquad (8.1) \\
&\le 2\Big(\int_{[a_2^-, a_2^- + \frac{2}{n}] \cup [a_2^+ - \frac{2}{n}, a_2^+]} \int_{Q_2} |\partial_2 u_1(x_1, t, x_3)|^2 \,\mathrm{d}(x_1, x_3)\,\mathrm{d}t\Big)^{1/2} \longrightarrow 0
\end{aligned}$$

as $n \to \infty$. Hence,

$$\partial_2 u_{1,n} = \rho_n^{(2)} * \big((\chi_n^{(2)})' \chi_n^{(3)} u_1\big) + \rho_n^{(2)} * \big(\chi_n^{(2)} \chi_n^{(3)} \partial_2 u_1\big) \longrightarrow \partial_2 u_1 \quad \text{in } L^2(Q)$$

as $n \to \infty$. The convergence

$$\partial_3 u_{1,n} \longrightarrow \partial_3 u_1 \quad \text{in } L^2(Q)$$

as $n \to \infty$ is seen similarly. This shows $u_{1,n} \to u_1$ in $H^1(Q)$ as $n \to \infty$. The functions u_2 and u_3 are treated in the same way.

Let Φ be the linear and bounded Stein extension operator that maps functions in $H^1(Q)$ to functions in $H^1(\mathbb{R}^3)$, see Theorem 5.24 in [1]. We extend v_1 by 0 outside of Q and set

$$v_{1,n,m} := \rho_n^{(2)} * \big(\rho_n^{(3)} * \Phi\big(\big(\rho_m^{(1)} * \big(\chi_m^{(1)} v_1\big)\big)|_Q\big)\big)|_Q$$

for all $n, m \geq n_0$. This function is in $H^1(Q)$ and it satisfies $\partial_3 v_{1,n,m} \in H^1(Q)$ and $v_{1,n,m} = 0$ on Γ_1, since the support of $\rho_m^{(1)} * \big(\chi_m^{(1)} v_1\big)$ has distance of at least $\frac{1}{2m}$ from Γ_1. Let $\eta > 0$. Using that $v_1 = 0$ on Γ_1, as in (8.1) one sees that

$$\big\|(\chi_m^{(1)})' v_1\big\|_{L^2} \longrightarrow 0$$

as $m \to \infty$. Letting the occurring derivatives acting on $\chi_m^{(1)} v_1$, we see that there exists an $\widetilde{m} = \widetilde{m}(\eta) \geq n_0$ such that

$$\Big\|\rho_{\widetilde{m}}^{(1)} * \big(\chi_{\widetilde{m}}^{(1)} v_1\big) - v_1\Big\|_{H^1} \leq \eta.$$

Furthermore, there exists an $\widetilde{n} = \widetilde{n}(\eta) \geq n_0$ such that

$$\Big\|v_{1,\widetilde{n},\widetilde{m}} - \Phi\big(\rho_{\widetilde{m}}^{(1)} * \big(\chi_{\widetilde{m}}^{(1)} v_1\big)\big)|_Q\Big\|_{H^1} \leq \eta,$$

so that

$$\|v_{1,\widetilde{n},\widetilde{m}} - v_1\|_{H^1} \leq \big(1 + \|\Phi\|_{\mathcal{B}(H^1(Q),H^1(\mathbb{R}^3))}\big)\eta.$$

Set $\widehat{v}_1 := v_{1,\widetilde{n},\widetilde{m}}$.

Because $\widehat{v}_1$ does not necessarily fulfil the first-order boundary condition on the forth component of elements of $D(A_Y)$, we define $\varphi_n := \chi_n^{(3)} \partial_3 \widehat{v}_1$ and

$$v_1^n(x) := v_1^n(x_1, x_2, x_3) := \widehat{v}_1(x_1, x_2, a_3^-) + \int_{a_3^-}^{x_3} \varphi_n(x_1, x_2, t)\,\mathrm{d}t$$

for almost all $(x_1, x_2) \in (a_1^-, a_1^+) \times (a_2^-, a_2^+)$. The trace of φ_n vanishes o Γ_3 due to $\chi_n^{(3)}$, so that $\partial_3 v_1^n = \varphi_n = 0$ on Γ_3. Lemma 7.10 shows that $v_1^n = 0$ on Γ_1 because $\widehat{v}_1 = 0$ on Γ_1. As a result, v_1^n also satisfies the other boundary condition $v_1^n = 0$ on Γ_1. From $\chi_n^{(3)} \in W^{1,\infty}(Q)$ we infer that φ_n belongs to $H^1(Q)$. The identity $\partial_3 v_1^n = \varphi_n$ now shows $v_1^n, \partial_3 v_1^n \in H^1(Q)$.

It remains to check that v_1^n converges to $\widehat{v}_1$ in $H^1(Q)$ as $n \to \infty$. Dominated convergence yields

$$\varphi_n - \partial_3 \widehat{v}_1 = \big(\chi_n^{(3)} - \mathbb{1}\big)\partial_3 \widehat{v}_1 \longrightarrow 0$$

in $L^2(Q)$ as $n \to \infty$. We thus obtain the limit

$$v_1^n(x_1, x_2, x_3) - \widehat{v}_1(x_1, x_2, x_3) = \int_{a_3^-}^{x_3} \big(\varphi_n(x_1, x_2, t) - \partial_3 \widehat{v}_1(x_1, x_2, t)\big)\,\mathrm{d}t \longrightarrow 0$$

in $L^2(Q)$ as $n \to \infty$ since

$$\begin{aligned}
&\left(\int_Q \left|\int_{a_3^-}^{x_3} \big(\varphi_n(x_1, x_2, t) - \partial_3 \widehat{v}_1(\widehat{x}_3, t)\big)\,\mathrm{d}t\right|^2 \mathrm{d}x\right)^{1/2} \\
&\quad \le \left(\int_Q (x_3 - a_3^-) \int_{a_3^-}^{x_3} \left|\varphi_n(x_1, x_2, t) - \partial_3 \widehat{v}_1(\widehat{x}_3, t)\right|^2 \mathrm{d}t\,\mathrm{d}x\right)^{1/2} \\
&\quad \le \left(\int_{a_3^-}^{a_3^+} (a_3^+ - a_3^-) \int_{a_1^-}^{a_1^+} \int_{a_2^-}^{a_2^+} \int_{a_3^-}^{a_3^+} \left|\varphi_n(x_1, x_2, t) - \partial_3 \widehat{v}_1(x_1, x_2, t)\right|^2 \right. \\
&\qquad\qquad \left. \mathrm{d}t\,\mathrm{d}x_2\,\mathrm{d}x_1\,\mathrm{d}x_3\right)^{1/2} \\
&\quad \le (a_3^+ - a_3^-) \cdot \|\varphi_n - \partial_3 \widehat{v}_1\|_{L^2} \longrightarrow 0
\end{aligned}$$

as $n \to \infty$. Moreover,

$$\partial_3 v_1^n(x_1, x_2, x_3) - \partial_3 \widehat{v}_1(x_1, x_2, x_3) = \varphi_n(x_1, x_2, x_3) - \partial_3 \widehat{v}_1(x_1, x_2, x_3) \longrightarrow 0$$

in $L^2(Q)$ as $n \to \infty$. Furthermore,

$$\partial_j \varphi_n - \partial_{3j} \widehat{v}_1 = \big(\chi_n^{(3)} - \mathbb{1}\big)\partial_{3j} \widehat{v}_1 \longrightarrow 0 \quad \text{for } j \in \{1, 2\}$$

in $L^2(Q)$ as $n \to \infty$, so that as above

$$\partial_j v_1^n(x_1, x_2, x_3) - \partial_j \widehat{v}_1(x_1, x_2, x_3) = \int_{a_3^-}^{x_3} \big(\partial_j \varphi_n(x_1, x_2, t) - \partial_{3j} \widehat{v}_1(x_1, x_2, t)\big)\,\mathrm{d}t \longrightarrow 0$$

in $L^2(Q)$ as $n \to \infty$ for $j \in \{1, 2\}$. This gives us $v_1^n \to \widehat{v}_1$ in $H^1(Q)$ as $n \to \infty$. Thus, we can choose an $\widehat{n} = \widehat{n}(\eta) \ge \widetilde{n}$ so large such that

$$\big\|v_1^{\widehat{n}} - v_1\big\|_{H^1} \le \big\|v_1^{\widehat{n}} - \widehat{v}_1\big\|_{H^1} + \|\widehat{v}_1 - v_1\|_{H^1} \le \big(2 + \|\Phi\|_{\mathcal{B}(H^1(Q), H^1(\mathbb{R}^3))}\big)\eta,$$

which shows the assertion. The components v_2 and v_3 are treated in the same way. □

We set

$$\kappa_Y := \frac{3\,\|\nabla\sigma\|_{L^\infty}}{4\delta} + \frac{3\,\|\sigma\|_{L^\infty}\,\|\nabla\varepsilon\|_{L^\infty}}{4\delta^2} + \frac{3\,\|\nabla\varepsilon\|_{L^\infty} + 3\,\|\nabla\mu\|_{L^\infty}}{2\delta^2}.$$

Lemma 8.5. *The operators $A_Y - \kappa_Y I$ and $B_Y - \kappa_Y I$ are dissipative on Y.*

Proof:
Let $(u,v) \in D(A_Y)$. With integration by parts we see

$$\begin{aligned}
&\sum_{j=1}^{3} \int_Q \big(\partial_j C_1 v \cdot \partial_j u + \partial_j C_2 u \cdot \partial_j v\big)\,\mathrm{d}x \\
&\quad = \sum_{j=1}^{3} \int_Q \big(\partial_{j2} v_3\, \partial_j u_1 + \partial_{j3} v_1\, \partial_j u_2 + \partial_{j1} v_2\, \partial_j u_3 \\
&\qquad + \partial_{j3} u_2\, \partial_j v_1 + \partial_{j1} u_3\, \partial_j v_2 + \partial_{j2} u_1\, \partial_j v_3\big)\,\mathrm{d}x \\
&\quad = 0,
\end{aligned} \tag{8.2}$$

where we have used the boundary properties of u from Corollary 8.2 and the ones of v from $D(A_Y)$ in Lemma 7.17 to get rid of the boundary integrals. Thus, we have together with (7.26), (7.28) and Young's inequality that

$$\begin{aligned}
&\mathrm{Re}\big(A(u,v) \mid (u,v)\big)_Y \\
&= \int_Q \Big(-\frac{\sigma\varepsilon}{2\varepsilon}\,|u|^2 + \frac{\varepsilon}{\varepsilon} C_1 v \cdot u + \frac{\mu}{\mu} C_2 u \cdot v - \varepsilon \sum_{j=1}^{3} \partial_j\big(\tfrac{\sigma}{2\varepsilon} u\big) \cdot \partial_j u \\
&\qquad + \varepsilon \sum_{j=1}^{3} \partial_j\big(\tfrac{1}{\varepsilon} C_1 v\big) \cdot \partial_j u + \mu \sum_{j=1}^{3} \partial_j\big(\tfrac{1}{\mu} C_2 u\big) \cdot \partial_j v\Big)\,\mathrm{d}x \\
&= -\int_Q \frac{\sigma}{2}\,|u|^2\,\mathrm{d}x - \int_Q \frac{\sigma}{2}\,|\partial u|^2\,\mathrm{d}x - \sum_{j=1}^{3} \int_Q \big(\tfrac{\partial_j \sigma}{2} - \tfrac{\sigma \partial_j \varepsilon}{2\varepsilon}\big) u \cdot \partial_j u\,\mathrm{d}x \\
&\quad - \sum_{j=1}^{3} \int_Q \tfrac{\partial_j \varepsilon}{\varepsilon} C_1 v \cdot \partial_j u\,\mathrm{d}x - \sum_{j=1}^{3} \int_Q \tfrac{\partial_j \mu}{\mu} C_2 u \cdot \partial_j v\,\mathrm{d}x \\
&\le \Big(\frac{\|\nabla\sigma\|_{L^\infty}}{4\delta} + \frac{\|\sigma\|_{L^\infty} \|\nabla\varepsilon\|_{L^\infty}}{4\delta^2}\Big) \int_Q \big(3\varepsilon\,|u|^2 + \varepsilon\,|\partial u|^2\big)\,\mathrm{d}x \\
&\quad + \frac{\|\nabla\varepsilon\|_{L^\infty} + \|\nabla\mu\|_{L^\infty}}{2\delta^2} \int_Q \big(3\varepsilon\,|\partial u|^2 + 3\mu\,|\partial v|^2\big)\,\mathrm{d}x \\
&\le \kappa_Y\, \|(u,v)\|_Y^2\,,
\end{aligned}$$

where $|\partial u|$ and $|\partial v|$ denote the Frobenius norm of the Jacobi matrix of u and v, respectively.

Let now $(u,v) \in D(B_Y)$. The identity (8.2) (with interchanged roles of u and v) yields together with (7.26), (7.28) and Young's inequality in the same way as above that

$$\mathrm{Re}\big(B(u,v) \mid (u,v)\big)_Y$$

$$
\begin{aligned}
&= \int_Q \Big(-\frac{\sigma\varepsilon}{2\varepsilon} |u|^2 - \frac{\varepsilon}{\varepsilon} C_2 v \cdot u - \frac{\mu}{\mu} C_1 u \cdot v - \varepsilon \sum_{j=1}^{3} \partial_j \big(\tfrac{\sigma}{2\varepsilon} u\big) \cdot \partial_j u \\
&\qquad - \varepsilon \sum_{j=1}^{3} \partial_j \big(\tfrac{1}{\varepsilon} C_2 v\big) \cdot \partial_j u - \mu \sum_{j=1}^{3} \partial_j \big(\tfrac{1}{\mu} C_1 u\big) \cdot \partial_j v \Big) \,\mathrm{d}x \\
&= - \int_Q \frac{\sigma}{2} |u|^2 \,\mathrm{d}x - \int_Q \frac{\sigma}{2} |\partial u|^2 \,\mathrm{d}x - \sum_{j=1}^{3} \int_Q \big(\tfrac{\partial_j \sigma}{2} - \tfrac{\sigma \partial_j \varepsilon}{2\varepsilon}\big) u \cdot \partial_j u \,\mathrm{d}x \\
&\quad - \sum_{j=1}^{3} \int_Q \tfrac{\partial_j \varepsilon}{\varepsilon} C_2 v \cdot \partial_j u \,\mathrm{d}x - \sum_{j=1}^{3} \int_Q \tfrac{\partial_j \mu}{\mu} C_1 u \cdot \partial_j v \,\mathrm{d}x \\
&\leq \Big(\frac{\|\nabla\sigma\|_{L^\infty}}{4\delta} + \frac{\|\sigma\|_{L^\infty} \|\nabla\varepsilon\|_{L^\infty}}{4\delta^2} \Big) \int_Q \big(3\varepsilon |u|^2 + \varepsilon |\partial u|^2 \big) \,\mathrm{d}x \\
&\quad + \frac{\|\nabla\varepsilon\|_{L^\infty} + \|\nabla\mu\|_{L^\infty}}{2\delta^2} \int_Q \big(3\varepsilon |\partial u|^2 + 3\mu |\partial v|^2 \big) \,\mathrm{d}x \\
&\leq \kappa_Y \, \|(u,v)\|_Y^2 \,,
\end{aligned}
$$

which finishes the proof. □

Lemma 8.6. *The operators $(1+\kappa_Y)I - A_Y$ and $(1+\kappa_Y)I - B_Y$ have dense range in Y.*

Proof:
We first deal with the operator $(1+\kappa_Y)I - A_Y$. Because we know from Lemma 8.4 that $D(A_Y)$ is dense in Y, it is sufficient to show that the range of $(1+\kappa_Y)I - A_Y$ contains $D(A_Y)$. Let $(f,g) \in D(A_Y)$. We look for fields $(u,v) \in D(A_Y)$ with $\big((1+\kappa_Y)I - A\big)(u,v) = (f,g)$, i.e.

$$
\big(1 + \kappa_Y + \tfrac{\sigma}{2\varepsilon}\big) u_1 - \frac{1}{\varepsilon} \partial_2 v_3 = f_1, \qquad (1+\kappa_Y) v_3 - \frac{1}{\mu} \partial_2 u_1 = g_3, \tag{8.3a}
$$

$$
\big(1 + \kappa_Y + \tfrac{\sigma}{2\varepsilon}\big) u_2 - \frac{1}{\varepsilon} \partial_3 v_1 = f_2, \qquad (1+\kappa_Y) v_1 - \frac{1}{\mu} \partial_3 u_2 = g_1, \tag{8.3b}
$$

$$
\big(1 + \kappa_Y + \tfrac{\sigma}{2\varepsilon}\big) u_3 - \frac{1}{\varepsilon} \partial_1 v_2 = f_3, \qquad (1+\kappa_Y) v_2 - \frac{1}{\mu} \partial_1 u_3 = g_2. \tag{8.3c}
$$

Plugging in each line the second equation into the first one, we get with the abbreviation $D_j := \partial_j \frac{1}{\mu} \partial_j$ for $j \in \{1,2,3\}$ the equations

$$
\big(\varepsilon(1+\kappa_Y) + \tfrac{\sigma}{2}\big) u_1 - \frac{1}{1+\kappa_Y} D_2 u_1 = \varepsilon f_1 + \frac{1}{1+\kappa_Y} \partial_2 g_3 =: h_1, \tag{8.4a}
$$

$$
\big(\varepsilon(1+\kappa_Y) + \tfrac{\sigma}{2}\big) u_2 - \frac{1}{1+\kappa_Y} D_3 u_2 = \varepsilon f_2 + \frac{1}{1+\kappa_Y} \partial_3 g_1 =: h_2, \tag{8.4b}
$$

$$
\big(\varepsilon(1+\kappa_Y) + \tfrac{\sigma}{2}\big) u_3 - \frac{1}{1+\kappa_Y} D_1 u_3 = \varepsilon f_3 + \frac{1}{1+\kappa_Y} \partial_1 g_2 =: h_3. \tag{8.4c}
$$

Let $j \in \{1,2,3\}$. Since $(f,g) \in D(A_Y)$, the function h_j belongs to $H^1(Q)$ and satisfies $h_j = 0$ on $\Gamma \setminus \Gamma_j$. We define

$$D(D_j) := \{\varphi \in L^2(Q) \mid \partial_j\varphi \in L^2(Q),\ \ D_j\varphi \in L^2(Q),\ \ \varphi = 0 \text{ on } \Gamma_j\}.$$

Using the general assumptions on μ, we obtain

$$D(D_j) = \{\varphi \in L^2(Q) \mid \partial_j\varphi \in L^2(Q),\ \ \partial_j^2\varphi \in L^2(Q),\ \ \varphi = 0 \text{ on } \Gamma_j\}.$$

Furthermore, we set

$$D(\partial_j) := \{\varphi \in L^2(Q) \mid \partial_j\varphi \in L^2(Q),\ \ \varphi = 0 \text{ on } \Gamma_j\}.$$

Let $j = 2$. We define the operator L by

$$Lw := \big((1+\kappa_Y)\varepsilon + \tfrac{\sigma}{2}\big)w - \frac{1}{1+\kappa_Y}\partial_2\big(\tfrac{1}{\mu}\partial_2 w\big)$$

for $w \in D(D_2)$. As in the proof of Lemma 4.3 in [37] we obtain a function w_1 in $D(D_2)$ with $Lw_1 = h_1$. Moreover, L is invertible. From $\partial_k\partial_2 w_1 \in H^{-1}(Q)$ and the general assumptions on μ, we infer $\frac{1}{\mu}\partial_2\partial_k w_1 \in H^{-1}(Q)$ and thus $D_2\partial_k w_1 \in H^{-2}(Q)$ for all $k \in \{1,2,3\}$. Let $\varphi \in H_0^2(Q)$ and $k \in \{1,2,3\}$. We can thus compute

$$\begin{aligned}
\langle L\partial_k w_1, \varphi\rangle_{H^{-2}\times H_0^2} &= \big\langle \partial_k w_1, \big((1+\kappa_Y)\varepsilon + \tfrac{\sigma}{2}\big)\varphi\big\rangle_{H^{-1}\times H_0^1} - \frac{1}{1+\kappa_Y}\Big\langle \partial_2\tfrac{1}{\mu}\partial_2\partial_k w_1, \varphi\Big\rangle_{H^{-2}\times H_0^2} \\
&= -\int_Q w_1\Big(\big(\partial_k\big((1+\kappa_Y)\varepsilon + \tfrac{\sigma}{2}\big)\big)\varphi + \big((1+\kappa_Y)\varepsilon + \tfrac{\sigma}{2}\big)\partial_k\varphi\Big)\,\mathrm{d}x \\
&\quad + \frac{1}{1+\kappa_Y}\Big\langle \partial_k\partial_2 w_1, \tfrac{1}{\mu}\partial_2\varphi\Big\rangle_{H^{-1}\times H_0^1} \\
&= -\int_Q \big((1+\kappa_Y)\varepsilon + \tfrac{\sigma}{2}\big)w_1\partial_k\varphi\,\mathrm{d}x - \int_Q \big(\partial_k\big((1+\kappa_Y)\varepsilon + \tfrac{\sigma}{2}\big)\big)w_1\varphi\,\mathrm{d}x \\
&\quad - \frac{1}{1+\kappa_Y}\int_Q (\partial_2 w_1)\big(\big(\partial_k\tfrac{1}{\mu}\big)\partial_2\varphi + \tfrac{1}{\mu}\partial_2\partial_k\varphi\big)\,\mathrm{d}x \\
&= -\int_Q (\partial_k\varphi)Lw_1\,\mathrm{d}x - \int_Q \big(\partial_k\big((1+\kappa_Y)\varepsilon + \tfrac{\sigma}{2}\big)\big)w_1\varphi\,\mathrm{d}x \\
&\quad - \frac{1}{1+\kappa_Y}\int_Q \big(\partial_k\tfrac{1}{\mu}\big)(\partial_2 w_1)\partial_2\varphi\,\mathrm{d}x \\
&= \int_Q (\partial_k h_1)\varphi\,\mathrm{d}x - \int_Q \big(\partial_k\big((1+\kappa_Y)\varepsilon + \tfrac{\sigma}{2}\big)\big)w_1\varphi\,\mathrm{d}x \\
&\quad + \frac{1}{1+\kappa_Y}\Big\langle \partial_2\big(\big(\partial_k\tfrac{1}{\mu}\big)\partial_2 w_1\big), \varphi\Big\rangle_{D(\partial_2)^*\times D(\partial_2)},
\end{aligned}$$

using that $H_0^2(Q) \hookrightarrow D(\partial_2)$. Note that the function $\partial_k\big(\frac{1}{\mu}\big)\partial_2 w_1$ belongs to $L^2(Q)^3$. By the density of $H^2(Q)$ in $D(\partial_2)$, this identity thus holds true for all $\varphi \in D(\partial_2)$ and

$$\begin{aligned}
L\partial_k w_1 &= \partial_k h_1 - \Big(\partial_k\big((1+\kappa_Y)\varepsilon + \tfrac{\sigma}{2}\big)\big)w_1 + \frac{1}{1+\kappa_Y}\partial_2\big(\big(\partial_k\tfrac{1}{\mu}\big)\partial_2 w_1\big)\Big) \\
&=: \psi_1(h_1) \in H^{-2}(Q).
\end{aligned} \tag{8.5}$$

We observe that the operator L is given by the symmetric, closed, positive definite and densely defined bilinear form

$$(w,\widetilde{w}) \mapsto \Big(\big((1+\kappa_Y)\varepsilon+\tfrac{\sigma}{2}\big)w,\widetilde{w}\Big)_{L^2} + \frac{1}{1+\kappa_Y}\Big(\tfrac{1}{\mu}\partial_2 w,\partial_2\widetilde{w}\Big)_{L^2}$$

on $D(\partial_2)$. Thus, L is self-adjoint due to the mentioned properties of the form by Proposition 1.24 in [61]. Theorem VI.2.23 in [46] yields the equivalence $D(\partial_2) \cong D(L^{1/2})$. Thus, $D(\partial_2)^* \cong D(L^{1/2})^*$, so that $\partial_k w_1 = L_{-1}^{-1}\psi_1(h_1) \in D(\partial_2) \cong D(L^{1/2})$. Here, L_{-1}^{-1} is the extension of L^{-1} to the Sobolev space of order -1, see Section II.5a in [23] and also Section 9.2. Because this is true for all $k \in \{1,2,3\}$, we have that $\partial_2 w_1$ is contained in $H^1(Q)$.

We now verify the boundary conditions for w_1. From $w_1 \in D(D_2)$ we know that $w_1 = 0$ on Γ_2. Moreover, we have $w_1 = L^{-1}h_1$ and as remarked above, $h_1 = 0$ on Γ_3. The proof of Lemma 8.4 shows that there exist functions $h_{1,n}$ in $H^1(Q)$ whose support has a distance of at least $\frac{1}{2n}$ to $\Gamma_2 \cup \Gamma_3$ and satisfy $h_{1,n} \to h_1$ in $H^1(Q)$ as $n \to \infty$. Set $w_{1,n} := L^{-1}h_{1,n} \in D(D_2)$ and take a function $\chi_n \in C_c^\infty((a_3^-,a_3^+))$ that is constant on $[a_3^- + \frac{1}{2n}, a_3^+ - \frac{1}{2n}]$. We then obtain

$$h_{1,n} = \chi_n h_{1,n} = \chi_n L w_{1,n} = L(\chi_n w_{1,n}).$$

Note that $\chi_n w_{1,n}$ belongs to $D(D_2)$ and that $w_{1,n} = \chi_n w_{1,n}$ vanishes on Γ_3. We know that $w_{1,n} = L^{-1}h_{1,n}$ tends to w_1 in $D(D_2)$. The above arguments further imply

$$\begin{aligned}
\|\partial_k(w_{1,n}-w_1)\|_{L^2} &= \big\|L_{-1}^{-1}\psi_1(h_{1,n}-h_1)\big\|_{L^2} \le c\,\|\psi_1(h_{1,n}-h_1)\|_{D(\partial_2)} \\
&\le c\Big(\|\partial_k h_{1,n}-\partial_k h_1\|_{L^2} + \Big\|\big(\partial_k\big((1+\kappa_Y)\varepsilon+\tfrac{\sigma}{2}\big)\big)(w_{1,n}-w_1)\Big\|_{L^2} \\
&\qquad + \frac{1}{1+\kappa_Y}\Big\|\big(\partial_k\big(\tfrac{1}{\mu}\big)\big)\partial_2(w_{1,n}-w_1)\Big\|_{L^2}\Big) \\
&\longrightarrow 0
\end{aligned}$$

as $n \to \infty$. Therefore, $w_{1,n}$ converges to w_1 in $H^1(Q)$ and so w_1 also vanishes on Γ_3 by the continuity of the trace.

We define

$$\widetilde{w}_3 := \frac{1}{1+\kappa_Y}g_3 + \frac{1}{1+\kappa_Y}\frac{1}{\mu}\partial_2 w_1 \in H^1(Q),$$

compare (8.3a). We then have $\widetilde{w}_3 = 0$ on Γ_3 since $g_3 = 0$ on Γ_3 by $(f,g) \in D(A_Y)$ and $\partial_2 w_1 = 0$ on Γ_3 with Lemma 7.10. In the above equation we take the derivative with respect to the second variable and plug in $Lw_1 = h_1$. It follows

$$\partial_2\widetilde{w}_3 = \frac{1}{1+\kappa_Y}\partial_2 g_3 + \big((1+\kappa_Y)\varepsilon+\tfrac{1}{2}\sigma\big)w_1 - h_1$$

in $L^2(Q)$. The definition of h_1 in (8.4a) then yields

$$\partial_2 \widetilde{w}_3 = -\varepsilon f_1 + \big(\varepsilon(1+\kappa_Y) + \tfrac{\sigma}{2}\big) w_1$$

in $L^2(Q)$. So, (8.3a) is valid for $u_1 := w_1$ and $v_1 := \widetilde{w}_3$. Due to the regularity of the right-hand side $\partial_2 \widetilde{w}_3$ belongs to $H^1(Q)$. The boundary condition $\partial_2 \widetilde{w}_3 = 0$ on Γ_2 now follows with Lemma 7.5 from $f_1 = 0$ and $w_1 = 0$ on Γ_2. The other components of w and $\widetilde{w}$ are treated in the same way. Altogether, we hence have $(w, \widetilde{w}) \in D(A_Y)$ and $A(w, \widetilde{w}) = (f, g)$.

For the operator B_Y we get

$$\begin{aligned}
\big(1+\kappa_Y+\tfrac{\sigma}{2\varepsilon}\big)u_1 + \frac{1}{\varepsilon}\partial_3 v_2 = f_1, \qquad & (1+\kappa_Y)v_2 + \frac{1}{\mu}\partial_3 u_1 = g_2,\\
\big(1+\kappa_Y+\tfrac{\sigma}{2\varepsilon}\big)u_2 + \frac{1}{\varepsilon}\partial_1 v_3 = f_2, \qquad & (1+\kappa_Y)v_3 + \frac{1}{\mu}\partial_1 u_2 = g_3,\\
\big(1+\kappa_Y+\tfrac{\sigma}{2\varepsilon}\big)u_3 + \frac{1}{\varepsilon}\partial_2 v_1 = f_3, \qquad & (1+\kappa_Y)v_1 + \frac{1}{\mu}\partial_2 u_3 = g_1
\end{aligned}$$

and

$$\begin{aligned}
\big(\varepsilon(1+\kappa_Y)+\tfrac{\sigma}{2}\big)u_1 - \frac{1}{1+\kappa_Y} D_3 u_1 &= \varepsilon f_1 - \frac{1}{1+\kappa_Y}\partial_3 g_2 =: h_1,\\
\big(\varepsilon(1+\kappa_Y)+\tfrac{\sigma}{2}\big)u_2 - \frac{1}{1+\kappa_Y} D_1 u_2 &= \varepsilon f_2 - \frac{1}{1+\kappa_Y}\partial_1 g_3 =: h_2,\\
\big(\varepsilon(1+\kappa_Y)+\tfrac{\sigma}{2}\big)u_3 - \frac{1}{1+\kappa_Y} D_2 u_3 &= \varepsilon f_3 - \frac{1}{1+\kappa_Y}\partial_2 g_1 =: h_3.
\end{aligned}$$

instead of (8.3) and (8.4). Now we get the statement for B_Y in the same way as the one for A_Y. □

Proposition 8.7. *(a) The operators A_Y and B_Y generate C_0-semigroups on Y whose norms are bounded by $e^{\kappa_Y t}$. The restrictions of $(I-\tau A)^{-1}$ and $(I-\tau B)^{-1}$ to Y are the operators $(I-\tau A_Y)^{-1}$ and $(I-\tau B_Y)^{-1}$, respectively. The semigroup estimate implies*

$$\big\|(I-\tau A_Y)^{-1}\big\|_{\mathcal{B}(Y)} \le \frac{1}{1-\tau\kappa_Y} \qquad \text{and} \qquad \big\|(I-\tau B_Y)^{-1}\big\|_{\mathcal{B}(Y)} \le \frac{1}{1-\tau\kappa_Y}$$

for all $0 < \tau < \frac{1}{\kappa_Y}$, which means in particular

$$\big\|(I-\tau A_Y)^{-1}\big\|_{\mathcal{B}(Y)} \le 2 \qquad \text{and} \qquad \big\|(I-\tau B_Y)^{-1}\big\|_{\mathcal{B}(Y)} \le 2$$

for all $0 < \tau \le \frac{1}{2\kappa_Y}$. Furthermore, the operators $A_Y - \kappa_Y I$ and $B_Y - \kappa_Y I$ are maximally dissipative on Y.

(b) *The parts of A_Y^* and B_Y^* of A^* and B^* in Y generate C_0-semigroups on Y whose norms are bounded by $e^{\kappa_Y t}$. The restrictions of $(I-\tau A^*)^{-1}$ and $(I-\tau B^*)^{-1}$ to Y are the operators $(I-\tau A_Y^*)^{-1}$ and $(I-\tau B_Y^*)^{-1}$, respectively. The semigroup estimate implies*

$$\left\|(I-\tau A_Y^*)^{-1}\right\|_{\mathcal{B}(Y)} \leq \frac{1}{1-\tau\kappa_Y} \qquad \text{and} \qquad \left\|(I-\tau B_Y^*)^{-1}\right\|_{\mathcal{B}(Y)} \leq \frac{1}{1-\tau\kappa_Y}$$

for all $0<\tau<\frac{1}{\kappa_Y}$, which means in particular

$$\left\|(I-\tau A_Y^*)^{-1}\right\|_{\mathcal{B}(Y)} \leq 2 \qquad \text{and} \qquad \left\|(I-\tau B_Y^*)^{-1}\right\|_{\mathcal{B}(Y)} \leq 2$$

for all $0<\tau\leq\frac{1}{2\kappa_Y}$. Furthermore, the operators $A_Y^*-\kappa_Y I$ and $B_Y^*-\kappa_Y I$ are maximally dissipative on Y.

(c) *We define the function*

$$\gamma_\tau(z) := \frac{1+\tau z}{1-\tau z}$$

on $\mathbb{C}\setminus\{\frac{1}{\tau}\}$. Then there exists a $\widetilde{\tau}\in\left(0,\frac{1}{\kappa_Y}\right)$ such that

$$\begin{aligned} &\|\gamma_\tau(A_Y)\|_{\mathcal{B}(Y)} \leq e^{3\kappa_Y\tau}, && \|\gamma_\tau(B_Y)\|_{\mathcal{B}(Y)} \leq e^{3\kappa_Y\tau},\\ &\|\gamma_\tau(A_Y^*)\|_{\mathcal{B}(Y)} \leq e^{3\kappa_Y\tau}, && \|\gamma_\tau(B_Y^*)\|_{\mathcal{B}(Y)} \leq e^{3\kappa_Y\tau} \end{aligned}$$

for all $0<\tau<\widetilde{\tau}$.

Proof:
Due to Lemma 7.14, the statements for the adjoint operators are seen as the other ones. Therefore, we only show the proofs for A_Y and B_Y.

(a) Due to the Theorem of Lumer–Phillips, Lemma 8.4, 8.5 and 8.6 imply that $A_Y-\kappa_Y I$ and $B_Y-\kappa_Y I$ generate contraction semigroups on Y. So $(0,\infty)$ is in the resolvent set of $A_Y-\kappa_Y I$ and $B_Y-\kappa_Y I$, which together with Lemma 8.5 yields the maximal dissipativity. Moreover, A_Y and B_Y generate C_0-semigroups $T_A(\cdot)$ and $T_B(\cdot)$ on Y with $\|T_A(t)\|_{\mathcal{B}(Y)}\leq e^{\kappa_Y t}$ and $\|T_B(t)\|_{\mathcal{B}(Y)}\leq e^{\kappa_Y t}$ for all $t\geq 0$. Let $0<\tau<\frac{1}{\kappa_Y}$, see Section II.2.2 in [23]. The statement of the restrictions of the resolvents follows since for complex numbers with a larger real part than the growth bound of a semigroup, the resolvent is the Laplace transform of the semigroup, see Theorem II.1.10 in [23]. This also gives

$$\begin{aligned} \left\|(I-\tau A_Y)^{-1}y\right\|_Y &= \left\|\tfrac{1}{\tau}\left(\tfrac{1}{\tau}I-A_Y\right)^{-1}y\right\|_Y \leq \frac{1}{\tau}\left\|\int_0^\infty e^{-\frac{1}{\tau}t}T_A(t)y\,\mathrm{d}t\right\|_Y\\ &\leq \frac{1}{\tau}\int_0^\infty e^{\kappa_Y t-\frac{1}{\tau}t}\|y\|_Y\,\mathrm{d}t = \frac{1}{\tau}\frac{1}{\frac{1}{\tau}-\kappa_Y}\|y\|_Y = \frac{1}{1-\tau\kappa_Y}\|y\|_Y \end{aligned}$$

for all $y\in Y$. The estimate for $(I-\tau B_Y)^{-1}$ is done in the same way.

(c) Let again $0 < \tau < \frac{1}{\kappa_Y}$. Due to $1 + \tau(z - \kappa_Y) \neq 0$ for $\operatorname{Re} z > 0$ we can define

$$\widetilde{\gamma}_\tau(z) := \frac{1 - \tau(z - \kappa_Y)}{1 + \tau(z - \kappa_Y)}$$

on $\{z \in \mathbb{C} \mid \operatorname{Re} z > 0\}$. We observe $\gamma_\tau(z) = \widetilde{\gamma}_\tau(\kappa_Y - z)$. For $r > 0$ we look at the mapping

$$s \mapsto \widetilde{\gamma}_\tau(r + is) = \frac{1 - \tau(r + is - \kappa_Y)}{1 + \tau(r + is - \kappa_Y)}$$

for $s \in \mathbb{R}$. Because $\widetilde{\gamma}_\tau$ is a Möbius transform, the generalized circles $\{r + is \mid s \in \mathbb{R}\}$ are mapped by $\widetilde{\gamma}_\tau$ on a generalized circles K_r, i.e. either a circle or a straight line. From $\widetilde{\gamma}_\tau(r + is) = \overline{\widetilde{\gamma}_\tau(r - is)}$ we conclude that the K_r are symmetric with respect to the real axis and from $\lim_{s\to\pm\infty} \widetilde{\gamma}_\tau(r + is) = -1$ we infer that K_r are circles through -1 and $\widetilde{\gamma}_\tau(r) \in \mathbb{R}$. Therefore, the point on the K_r with the largest distance to the origin is either -1 or $\widetilde{\gamma}_\tau(r)$. From $\widetilde{\gamma}_\tau(0) = \frac{1+\tau\kappa_Y}{1-\tau\kappa_Y} > 1$, $\lim_{r\to\infty} \widetilde{\gamma}_\tau(r) = -1$ and $\widetilde{\gamma}'_\tau(r) = -\frac{2\tau}{(1+\tau(r-\kappa_Y))^2} < 0$ for all $r \in (0, \infty)$ infer

$$\begin{aligned}\sup_{\operatorname{Re} z>0} \|\widetilde{\gamma}_\tau(z)\| &= \sup_{r>0} \|\widetilde{\gamma}_\tau(r + \cdot)\|_\infty = \sup_{r>0} \max\{1, |\widetilde{\gamma}_\tau(r)|\} \\ &= \max\Big\{1, \sup_{r>0} |\widetilde{\gamma}_\tau(r)|\Big\} = \widetilde{\gamma}_\tau(0) = \frac{1 + \tau\kappa_Y}{1 - \tau\kappa_Y}.\end{aligned}$$

We define

$$\phi(\tau) := \ln\Big(\frac{1 + \tau\kappa_Y}{1 - \tau\kappa_Y}\Big)$$

for $\tau \in (0, \frac{1}{\kappa_Y})$ and $\phi(0) := 0$ and see that ϕ is continuous on $\big[0, \frac{1}{\kappa_Y}\big)$. By applying L'Hospital's rule we get

$$\phi'(0) = \lim_{\tau\to 0} \frac{\phi(\tau)}{\tau} = \lim_{\tau\to 0} \frac{\frac{2\kappa_Y}{(1-\tau\kappa_Y)^2}}{\frac{1+\tau\kappa_Y}{1-\tau\kappa_Y}} = \lim_{\tau\to 0} \frac{2\kappa_Y}{1 - \tau^2\kappa_Y^2} = 2\kappa_Y$$

and hence have with $\phi'(\tau) = \frac{2\kappa_Y}{1-\tau^2\kappa_Y^2}$ for $\tau > 0$ that $\phi \in C^1([0, \frac{1}{\kappa_Y}))$. So, there exists a $\widetilde{\tau} \in (0, \frac{1}{\kappa_Y})$ with $\phi(\tau) \leq 3\kappa_Y\tau$ and therefore $\sup_{\operatorname{Re} z>0} \|\widetilde{\gamma}_\tau(z)\| \leq e^{3\kappa_Y\tau}$ for all $\tau \in (0, \widetilde{\tau})$. Because the operator $\kappa_Y I - A_Y = -(A_Y - \kappa_Y I)$ is maximal accretive by part (a), we can apply Theorem 11.5 of [48] and get a H^∞-functional calculus for $\kappa_Y I - A_Y$ together with the estimate

$$\|\gamma_\tau(A_Y)\|_{\mathcal{B}(Y)} = \|\widetilde{\gamma}_\tau(\kappa_Y I - A_Y)\|_{\mathcal{B}(Y)} \leq \sup_{\operatorname{Re} z>0} |\widetilde{\gamma}_\tau(z)| \leq e^{3\kappa_Y\tau}$$

for all $\tau \in (0, \widetilde{\tau})$. The other estimate is shown in the same way and the operators B_Y and B_Y^* are treated analogously. □

8.3. Properties of the splitting operators in the H^2-setting

We first use Lemma 7.10 to deduce from the definition of $D(A_Z)$ and $D(B_Z)$ further trace properties of these domains, see Subsection 7.2.2 for the definition of the operators A_Z and B_Z, as well as A_Y and B_Y. In addition, we still have those of Corollary 8.2 since $A_Z \subseteq A_Y$ and $B_Z \subseteq B_Y$.

Corollary 8.8. *(a) Let $(u,v) \in D(A_Z)$. Then*

$$\begin{aligned}
&\partial_2 u_2 = \partial_3 u_2 = \partial_2 u_3 = \partial_3 u_3 = \partial_2 v_1 = \partial_3 v_1 = 0 && \text{on } \Gamma_1,\\
&\partial_{23} u_2 = \partial_{33} u_2 = \partial_{22} u_3 = \partial_{23} u_3 = \partial_{33} u_3 = 0 && \text{on } \Gamma_1,\\
&\partial_{23} v_1 = \partial_{33} v_1 = \partial_{12} v_2 = \partial_{13} v_2 = 0 && \text{on } \Gamma_1,\\
&\partial_3 u_1 = \partial_1 u_1 = \partial_1 u_3 = \partial_3 u_3 = \partial_1 v_2 = \partial_3 v_2 = 0 && \text{on } \Gamma_2,\\
&\partial_{11} u_3 = \partial_{13} u_3 = \partial_{11} u_1 = \partial_{13} u_1 = \partial_{33} u_1 = 0 && \text{on } \Gamma_2,\\
&\partial_{12} v_2 = \partial_{13} v_2 = \partial_{12} v_3 = \partial_{23} v_3 = 0 && \text{on } \Gamma_2,\\
&\partial_1 u_1 = \partial_2 u_2 = \partial_1 u_2 = \partial_2 u_2 = \partial_1 v_3 = \partial_2 v_3 = 0 && \text{on } \Gamma_3,\\
&\partial_{12} u_1 = \partial_{22} u_1 = \partial_{11} u_2 = \partial_{12} u_2 = \partial_{22} u_2 = 0 && \text{on } \Gamma_3,\\
&\partial_{12} v_3 = \partial_{22} v_3 = \partial_{13} v_1 = \partial_{23} v_1 = 0 && \text{on } \Gamma_3.
\end{aligned}$$

(b) Let $(u,v) \in D(B_Z)$. Then

$$\begin{aligned}
&\partial_2 u_2 = \partial_3 u_2 = \partial_2 u_3 = \partial_3 u_3 = \partial_2 v_1 = \partial_3 v_1 = 0 && \text{on } \Gamma_1,\\
&\partial_{22} u_3 = \partial_{23} u_3 = \partial_{22} u_2 = \partial_{23} u_2 = \partial_{33} u_2 = 0 && \text{on } \Gamma_1,\\
&\partial_{22} v_1 = \partial_{23} v_1 = \partial_{12} v_3 = \partial_{13} v_3 = 0 && \text{on } \Gamma_1,\\
&\partial_3 u_1 = \partial_1 u_1 = \partial_1 u_3 = \partial_3 u_3 = \partial_1 v_2 = \partial_3 v_2 = 0 && \text{on } \Gamma_2,\\
&\partial_{13} u_1 = \partial_{33} u_1 = \partial_{11} u_3 = \partial_{13} u_3 = \partial_{33} u_3 = 0 && \text{on } \Gamma_2,\\
&\partial_{13} v_2 = \partial_{33} v_2 = \partial_{12} v_1 = \partial_{23} v_1 = 0 && \text{on } \Gamma_2,\\
&\partial_1 u_1 = \partial_2 u_2 = \partial_1 u_2 = \partial_2 u_2 = \partial_1 v_3 = \partial_2 v_3 = 0 && \text{on } \Gamma_3,\\
&\partial_{11} u_2 = \partial_{12} u_2 = \partial_{11} u_1 = \partial_{12} u_1 = \partial_{22} u_1 = 0 && \text{on } \Gamma_3,\\
&\partial_{11} v_3 = \partial_{12} v_3 = \partial_{13} v_2 = \partial_{23} v_2 = 0 && \text{on } \Gamma_3.
\end{aligned}$$

In the next lemmas we collect some properties of A_Z and B_Z.

Lemma 8.9. *Let $\varepsilon, \mu, \sigma \in W^{2,3}(Q)$. Then the operators A_Z and B_Z are closed in Z and densely defined on Z.*

PROOF:
1) To show the closedness of A_Z let $(u_n, v_n)_{n\in\mathbb{N}} \subseteq D(A_Z)$ be a sequence with $(u_n, v_n) \to (u, v)$ in Z and $A(u_n, v_n) \to (f, g)$ in Z as $n \to \infty$. The fields (u, v) satisfy the boundary conditions of Z by the continuity of the traces. Additionally, C_2u_n and C_1v_n converge to C_2u and C_1v in $H^2(Q)^3$, respectively, as $n \to \infty$ due to the assumptions on ε, μ and σ, and $A(u, v) = (f, g)$. Furthermore, we deduce $\frac{\sigma}{2}u_n \to \frac{\sigma}{2}u$ in $H^2(Q)^3$ and

$$C_1v_n = \left(-\tfrac{\sigma}{2}u_n + C_1v_n\right) + \tfrac{\sigma}{2}u_n \longrightarrow \varepsilon f + \tfrac{\sigma}{2}u$$

and $C_2u_n \to \mu g$ in $H^2(Q)^3$ as $n \to \infty$. So, C_2u and C_1v belong to $H^2(Q)^3$ and (u, v) satisfies the second-order boundary conditions of $D(A_Z)$. Altogether we have $(u, v) \in D(A_Z)$ and $A(u, v) = (f, g)$.

To show the closedness of B_Z let $(u_n, v_n)_{n\in\mathbb{N}} \subseteq D(B_Z)$ be a sequence with $(u_n, v_n) \to (u, v)$ in Z and $B(u_n, v_n) \to (f, g)$ in Z as $n \to \infty$. Then (u, v) satisfies the boundary conditions of Z. Moreover, C_1u_n and C_2v_n converge to C_1u and C_2v in $H^2(Q)^3$, respectively, as $n \to \infty$ again due to the assumptions on the coefficients, and $B(u, v) = (f, g)$. From $\varepsilon, \sigma \in W^{1,\infty}(Q) \cap W^{2,3}(Q)$ and $\varepsilon \geq \delta > 0$ we deduce $\frac{\sigma}{2}u_n \to \frac{\sigma}{2}u$ in $H^2(Q)^3$ and

$$C_2v_n = \left(-\tfrac{\sigma}{2}u_n + C_2v_n\right) + \tfrac{\sigma}{2}u_n \longrightarrow \varepsilon f + \tfrac{\sigma}{2}u$$

and $C_1u_n \to \mu g$ in $H^2(Q)^3$ as $n \to \infty$. So, C_1u and C_2v belong to $H^2(Q)^3$ and (u, v) fulfils the seond-order boundary conditions of $D(B_Z)$. Altogether we have $(u, v) \in D(B_Z)$ and $B(u, v) = (f, g)$.

2) Let $(u, v) \in Z$ and choose $n_0 \in \mathbb{N}$ with $n_0 \geq \frac{4}{d_{\min}}$. Let $(\rho_n^{(j)})_{n\in\mathbb{N}}$ be the standard sequences of symmetric C^∞-mollifiers with $\operatorname{supp}(\rho_n^{(j)}) \subseteq \left[-\frac{1}{n}, \frac{1}{n}\right]$ acting on the j-th variable. We define the cuboids

$$\begin{aligned} Q^{(1)} &:= (a_1^-, a_1^+) \times (2a_2^- - a_2^+, a_2^-) \times (a_3^-, a_3^+), \\ Q^{(2)} &:= (a_1^-, a_1^+) \times (a_2^+, 2a_2^+ - a_2^-) \times (a_3^-, a_3^+) \\ \text{and} \qquad \widetilde{Q} &:= (a_1^-, a_1^+) \times (2a_2^- - a_2^+, 2a_2^+ - a_2^-) \times (a_3^-, a_3^+), \end{aligned}$$

and extend u_1 in an antisymmetric way to $\widetilde{Q}$ by

$$\widetilde{u}_1(x_1, x_2, x_3) := \begin{cases} -u_1(x_1, 2a_2^- - x_2, x_3), & x_2 \in (2a_2^- - a_2^+, a_2^-), \\ u_1(x_1, x_2, x_3), & x_2 \in [a_2^-, a_2^+], \\ -u_1(x_1, 2a_2^+ - x_2, x_3), & x_2 \in (a_2^+, 2a_2^+ - a_2). \end{cases}$$

We first show that $\widetilde{u}_1$ is contained in $H^2(\widetilde{Q})$. Due to the regularity and the integrability of u_1 we only have to prove $\partial_2\widetilde{u}_1 \in L^2(\widetilde{Q})$ and $\partial_{22}\widetilde{u}_1 \in L^2(\widetilde{Q})$. Let $\varphi \in C_c^\infty(\widetilde{Q})$. Using $u_1 = 0$ on Γ_2, we have

$$\int_{\widetilde{Q}} \widetilde{u}_1\partial_2\varphi \,\mathrm{d}x = \int_{Q^{(1)}} \widetilde{u}_1\partial_2\varphi \,\mathrm{d}x + \int_{Q} \widetilde{u}_1\partial_2\varphi \,\mathrm{d}x + \int_{Q^{(2)}} \widetilde{u}_1\partial_2\varphi \,\mathrm{d}x$$

$$
\begin{aligned}
&= -\int_{Q^{(1)}} \varphi(x)(\partial_2 u_1)(x_1, 2a_2^- - x_2, x_3)\,\mathrm{d}x + \big[-u_1(x_1, 2a_2^- - x_2, x_3)\varphi(x)\big]_{\Gamma_2^-} \\
&\quad - \int_{Q} \varphi(x)(\partial_2 u_1)(x)\,\mathrm{d}x - \big[u_1(x)\varphi(x)\big]_{\Gamma_2^-} + \big[u_1(x)\varphi(x)\big]_{\Gamma_2^+} \\
&\quad - \int_{Q^{(2)}} \varphi(x)(\partial_2 u_1)(x_1, 2a_2^+ - x_2, x_3)\,\mathrm{d}x - \big[-u_1(x_1, 2a_2^+ - x_2, x_3)\varphi(x)\big]_{\Gamma_2^+} \\
&= -\int_{Q^{(1)}} \varphi(x)(\partial_2 u_1)(x_1, 2a_2^- - x_2, x_3)\,\mathrm{d}x - \int_{Q} \varphi(x)(\partial_2 u_1)(x_1, x_2, x_3)\,\mathrm{d}x \\
&\quad - \int_{Q^{(2)}} \varphi(x)(\partial_2 u_1)(x_1, 2a_2^+ - x_2, x_3)\,\mathrm{d}x.
\end{aligned}
$$

This shows that

$$
(\partial_2 \widetilde{u}_1)(x_1, x_2, x_3) = \begin{cases} (\partial_2 u_1)(x_1, 2a_2^- - x_2, x_3), & x_2 \in (2a_2^- - a_2^+, a_2^-), \\ (\partial_2 u_1)(x_1, x_2, x_3), & x_2 \in [a_2^- a_2^+], \\ (\partial_2 u_1)(x_1, 2a_2^+ - x_2, x_3), & x_2 \in (a_2^+, 2a_2^+ - a_2^-), \end{cases}
$$

is contained in $L^2(\widetilde{Q})$. Furthermore,

$$
\begin{aligned}
\int_{\widetilde{Q}} (\partial_2 \widetilde{u}_1)\partial_2 \varphi\,\mathrm{d}x &= \int_{Q^{(1)}} (\partial_2 \widetilde{u}_1)\partial_2 \varphi\,\mathrm{d}x + \int_{Q} (\partial_2 \widetilde{u}_1)\partial_2 \varphi\,\mathrm{d}x + \int_{Q^{(2)}} \partial_2 \widetilde{u}_1 \partial_2 \varphi\,\mathrm{d}x \\
&= -\int_{Q^{(1)}} -\varphi(x)(\partial_{22} u_1)(x_1, 2a_2^- - x_2, x_3)\,\mathrm{d}x \\
&\quad + \big[(\partial_2 u_1)(x_1, 2a_2^- - x_2, x_3)\varphi(x)\big]_{\Gamma_2^-} \\
&\quad - \int_{Q} \varphi(x)(\partial_{22} u_1)(x)\,\mathrm{d}x - \big[(\partial_2 u_1)(x)\varphi(x)\big]_{\Gamma_2^-} + \big[(\partial_2 u_1)(x)\varphi(x)\big]_{\Gamma_2^+} \\
&\quad - \int_{Q^{(2)}} -\varphi(x)(\partial_{22} u_1)(x_1, 2a_2^+ - x_2, x_3)\,\mathrm{d}x \\
&\quad - \big[(\partial_2 u_1)(x_1, 2a_2^+ - x_2, x_3)\varphi(x)\big]_{\Gamma_2^+} \\
&= -\int_{Q^{(1)}} -\varphi(x)(\partial_{22} u_1)(x_1, 2a_2^- - x_2, x_3)\,\mathrm{d}x \\
&\quad - \int_{Q} \varphi(x)(\partial_{22} u_1)(x_1, x_2, x_3)\,\mathrm{d}x \\
&\quad - \int_{Q^{(2)}} -\varphi(x)(\partial_{22} u_1)(x_1, 2a_2^+ - x_2, x_3)\,\mathrm{d}x,
\end{aligned}
$$

so that

$$
(\partial_{22} \widetilde{u}_1)(x_1, x_2, x_3) = \begin{cases} -(\partial_{22} u_1)(x_1, 2a_2^- - x_2, x_3), & x \in (2a_2^- - a_2^+, a_2^-), \\ (\partial_{22} u_1)(x_1, x_2, x_3), & x_2 \in [a_2^- a_2^+], \\ -(\partial_{22} u_1)(x_1, 2a_2^+ - x_2, x_3), & x_2 \in (a_2^+, 2a_2^+ - a_2^-), \end{cases}
$$

belongs to $L^2(\widetilde{Q})$.

Moreover, $\widetilde{u}_1 = 0$ on $\Gamma_2 \cup \Gamma_3$ and $\partial_1 \widetilde{u}_1 = 0$ on Γ_1 due to the definition of Z. Let $n \geq n_0$, extend $\widetilde{u}_1$ by 0 outside of $\widetilde{Q}$ and set

$$\widetilde{u}_1^n := (\rho_n^{(2)} * \widetilde{u}_1)|_{\widetilde{Q}}.$$

Then $\widetilde{u}_1^n$ and $\partial_2^k \widetilde{u}_1^n$ belong to $H^2(\widetilde{Q})$ for all $k \in \mathbb{N}$. Lemma 7.10 says that $\widetilde{u}_1^n = 0$ on Γ_3 and that $\partial_1 \widetilde{u}_1^n = \rho_n^{(2)} * \partial_1 \widetilde{u}_1 = 0$ on Γ_1. Additionally,

$$\begin{aligned}
&\widetilde{u}_1^n(x_1, a_2^-, x_3) \\
&\quad = \int_{-1/n}^{1/n} \rho_n^{(2)}(t) \widetilde{u}_1(x_1, a_2^- - t, x_3)\,\mathrm{d}t \\
&\quad = \int_{-1/n}^{0} \rho_n^{(2)}(t) u_1(x_1, a_2^- - t, x_3)\,\mathrm{d}t - \int_0^{1/n} \rho_n^{(2)}(t) u_1(x_1, a_2^- + t, x_3)\,\mathrm{d}t \\
&\quad = \int_0^{1/n} \rho_n^{(2)}(-s) u_1(x_1, a_2^- + s, x_3)\,\mathrm{d}s - \int_0^{1/n} \rho_n^{(2)}(t) u_1(x_1, a_2^- + t, x_3)\,\mathrm{d}t \\
&\quad = 0
\end{aligned}$$

for almost all $(x_1, x_3) \in (a_1^-, a_1^+) \times (a_3^-, a_3^+)$ due to the support and the symmetry of $\rho_n^{(2)}$. With the analogous calculation for a_2^+ instead of a_2^- we infer $\widetilde{u}_1^n = 0$ on Γ_2. Furthermore, we have $\widetilde{u}_1^n \to \widetilde{u}_1$ in $H^2(\widetilde{Q})$ as $n \to \infty$, so that $\widetilde{u}_1^n|_Q \to u_1$ in $H^2(Q)$ as $n \to \infty$. Let $\eta > 0$ and choose an $\widetilde{n} = \widetilde{n}(\eta) \geq n_0$ such that $\left\|\widetilde{u}_1^{\widetilde{n}} - u_1\right\|_{H^2} \leq \eta$ and set $\widehat{u}_1 := \widetilde{u}_1^{\widetilde{n}}$.

Because $\widehat{u}_1$ does not necessarily satisfy the second-order boundary condition of the first component of elements of $D(A_Z)$, we have to modify it once more. Let $\alpha, \beta \in C^\infty([a_2^-, a_2^+], [0,1])$ with $\alpha = 1$ on $\left[a_2^-, a_2^- + \frac{1}{3}(a_2^+ - a_2^-)\right]$, $\beta = 1$ on $\left[a_2^- + \frac{2}{3}(a_2^+ - a_2^-), a_2^+\right]$ and $\alpha + \beta = 1$ on $[a_2^-, a_2^+]$. So, $\alpha = 0$ on $\left[a_2^- + \frac{2}{3}(a_2^+ - a_2^-), a_2^+\right]$ and $\beta = 0$ on $\left[a_2^-, a_2^- + \frac{1}{3}(a_2^+ - a_2^-)\right]$. We set

$$\begin{aligned}
u_1^n(x_1, x_2, x_3) :&= \alpha(x_2)(x_2 - a_2^-)\partial_2 \widehat{u}_1(x_1, a_2^-, x_3) \\
&\quad + \alpha(x_2) \int_{a_2^-}^{x_2} \int_{a_2^-}^{t} \chi_n^{(2)}(s) \partial_{22} \widehat{u}_1(x_1, s, x_3)\,\mathrm{d}s\,\mathrm{d}t \\
&\quad + \beta(x_2)(x_2 - a_2^+)\partial_2 \widehat{u}_1(x_1, a_2^+, x_3) \\
&\quad + \beta(x_2) \int_{x_2}^{a_2^+} \int_{t}^{a_2^+} \chi_n^{(2)}(s) \partial_{22} \widehat{u}_1(x_1, s, x_3)\,\mathrm{d}s\,\mathrm{d}t,
\end{aligned}$$

where $\chi_n^{(2)}$ are the cut-off functions form the proof of Lemma 7.10 extended to $\mathbb{R}$ by zero. Then u_1^n and $\partial_2 u_1^n$ are contained in $H^2(Q)$ due to the regularity of $\widehat{u}_1$. Moreover, we have $u_1^n = 0$ on Γ_2 by the cut-off function in the definition of u_1^n and the supports of α and β. The trace condition $u_1^n = 0$ on Γ_3 follows from the boundary and regularity properties of

$\widehat{u}_1$ and Lemma 7.10. Furthermore, we have $\partial_1 u_1^n = 0$ on Γ_1 due to again the boundary and regularity properties of $\widehat{u}_1$ and Lemma 7.10. On Γ_2 we obtain

$$\partial_{22} u_1^n(x_1, x_2, x_3) = \chi_n^{(2)}(x_2)\partial_{22}\widehat{u}_1(x_1, x_2, x_3) = 0$$

due to the cut-off function. We use $\widehat{u}_1 = 0$ on Γ_2 to gain the representation

$$\begin{aligned}\widehat{u}_1(x_1, x_2, x_3) &= \alpha(x_2)(x_2 - a_2^-)\partial_2\widehat{u}_1(x_1, a_2^-, x_3) + \alpha(x_2)\int_{a_2^-}^{x_2}\int_{a_2^-}^{t}\partial_{22}\widehat{u}_1(x_1, s, x_3)\,\mathrm{d}s\,\mathrm{d}t \\ &\quad + \beta(x_2)(a_2^+ - x_2)\partial_2\widehat{u}_1(x_1, a_2^-, x_3) + \beta(x_2)\int_{x_2}^{a_2^+}\int_{t}^{a_2^+}\partial_{22}\widehat{u}_1(x_1, s, x_3)\,\mathrm{d}s\,\mathrm{d}t.\end{aligned}$$

Thus, it follows

$$\begin{aligned}u_1^n(x_1, x_2, x_3) - \widehat{u}_1(x_1, x_2, x_3) &= \alpha(x_2)\int_{a_2^-}^{x_2}\int_{a_2^-}^{t}\big(\chi_n^{(2)}(s) - 1\big)\partial_{22}\widehat{u}_1(x_1, s, x_3)\,\mathrm{d}s\,\mathrm{d}t \\ &\quad + \beta(x_2)\int_{x_2}^{a_2^+}\int_{t}^{a_2^+}\big(\chi_n^{(2)}(s) - 1\big)\partial_{22}\widehat{u}_1(x_1, s, x_3)\,\mathrm{d}s\,\mathrm{d}t.\end{aligned}$$

The theorem of dominated convergence then yields that

$$\|\partial_{jk}(u_1^n - \widehat{u}_1)\|_{L^2} \le c(a_2^+ - a_2^-)^2\,\big\|\big(\chi_n^{(2)} - 1\big)\partial_{22jk}\widehat{u}_1\big\|_{L^2} \longrightarrow 0$$

as $n \to \infty$ for all $j, k \in \{1, 3\}$. In the same way we see $u_1^n \to \widehat{u}_1$ and $\partial_j u_1^n \to \partial_j \widehat{u}_1$ in $L^2(Q)$ as $n \to \infty$ for all $j \in \{1, 3\}$. We treat the terms

$$\begin{aligned}\partial_2(u_1^n - \widehat{u}_1)(x_1, x_2, x_3) &= \alpha(x_2)\int_{a_2^-}^{x_2}\big(\chi_n^{(2)}(s) - 1\big)\partial_{22}\widehat{u}_1(x_1, s, x_3)\,\mathrm{d}s \\ &\quad + \alpha'(x_2)\int_{a_2^-}^{x_2}\int_{a_2^-}^{t}\big(\chi_n^{(2)}(s) - 1\big)\partial_{22}\widehat{u}_1(x_1, s, x_3)\,\mathrm{d}s \\ &\quad - \beta(x_2)\int_{x_2}^{a_2^+}\big(\chi_n^{(2)}(s) - 1\big)\partial_{22}\widehat{u}_1(x_1, s, x_3)\,\mathrm{d}s \\ &\quad + \beta'(x_2)\int_{x_2}^{a_2^+}\int_{t}^{a_2^+}\big(\chi_n^{(2)}(s) - 1\big)\partial_{22}\widehat{u}_1(x_1, s, x_3)\,\mathrm{d}s, \\ \partial_{j2}(u_1^n - \widehat{u}_1)(x_1, x_2, x_3) &= \alpha(x_2)\int_{a_2^-}^{x_2}\big(\chi_n^{(2)}(s) - 1\big)\partial_{22j}\widehat{u}_1(x_1, s, x_3)\,\mathrm{d}s \\ &\quad + \alpha'(x_2)\int_{a_2^-}^{x_2}\int_{a_2^-}^{t}\big(\chi_n^{(2)}(s) - 1\big)\partial_{22j}\widehat{u}_1(x_1, s, x_3)\,\mathrm{d}s \\ &\quad - \beta(x_2)\int_{x_2}^{a_2^+}\big(\chi_n^{(2)}(s) - 1\big)\partial_{22j}\widehat{u}_1(x_1, s, x_3)\,\mathrm{d}s \\ &\quad + \beta'(x_2)\int_{x_2}^{a_2^+}\int_{t}^{a_2^+}\big(\chi_n^{(2)}(s) - 1\big)\partial_{22j}\widehat{u}_1(x_1, s, x_3)\,\mathrm{d}s,\end{aligned}$$

$$\begin{aligned}\partial_{22}(u_1^n-\widehat{u}_1)(x_1,x_2,x_3) &= \big(\chi_n^{(2)}(x_2)-1\big)\partial_{22}\widehat{u}_1(x_1,x_2,x_3)\\ &\quad+\alpha'(x_2)\int_{a_2^-}^{x_2}\big(\chi_n^{(2)}(s)-1\big)\partial_{22}\widehat{u}_1(x_1,s,x_3)\,\mathrm{d}s\\ &\quad-\beta'(x_2)\int_{x_2}^{a_2^+}\big(\chi_n^{(2)}(s)-1\big)\partial_{22}\widehat{u}_1(x_1,s,x_3)\,\mathrm{d}s\\ &\quad+\alpha''(x_2)\int_{a_2^-}^{x_2}\int_{a_2^-}^{t}\big(\chi_n^{(2)}(s)-1\big)\partial_{22}\widehat{u}_1(x_1,s,x_3)\,\mathrm{d}s\,\mathrm{d}t\\ &\quad+\beta''(x_2)\int_{x_2}^{a_2^+}\int_{t}^{a_2^+}\big(\chi_n^{(2)}(s)-1\big)\partial_{22}\widehat{u}_1(x_1,s,x_3)\,\mathrm{d}s\,\mathrm{d}t\end{aligned}$$

in the same way. As a result, u_1^n tends to $\widehat{u}_1$ in $H^2(Q)$ as $n\to\infty$. Altogether we see that $u_1^n\to\widehat{u}_1$ in $H^2(Q)$ as $n\to\infty$. Thus, choosing $\widehat{n}\geq\widetilde{n}$ large enough, we have

$$\big\|u_1^{\widehat{n}}-u_1\big\|_{H^2}\leq\big\|u_1^{\widehat{n}}-\widehat{u}_1\big\|_{H^2}+\|\widehat{u}_1-u_1\|_{H^2}\leq 2\eta.$$

To deal with v_1, we redefine the cuboids from above to be

$$\begin{aligned}Q^{(1)} &:= (a_1^-,a_1^+)\times(a_2^-,a_2^+)\times(2a_3^--a_3^+,a_3^-),\\ Q^{(2)} &:= (a_1^-,a_1^+)\times(a_2^-,a_2^+)\times(a_3^+,2a_3^+-a_3^-)\\ \text{and}\qquad \widetilde{Q} &:= (a_1^-,a_1^+)\times(a_2^-,a_2^+)\times(2a_3^--a_3^+,2a_3^+-a_3^-),\end{aligned}$$

and extend v_1 in a symmetric way to $\widetilde{Q}$ by

$$\widetilde{v}_1(x_1,x_2,x_3):=\begin{cases}v_1(x_1,x_2,2a_3^--x_3), & x_3\in(2a_3^--a_3^+,a_3^-),\\ v_1(x_1,x_2,x_3), & x_3\in[a_3^-,a_3^+],\\ v_1(x_1,x_2,2a_3^+-x_3), & x_3\in(2a_3^+-a_3^-,a_3^+).\end{cases}$$

As, above, we first show $\widetilde{v}_1$ is contained in $H^2(\widetilde{Q})$. Due to the regularity and the integrability of v_1 we only have to prove $\partial_3\widetilde{v}_1\in L^2(\widetilde{Q})$ and $\partial_{33}\widetilde{v}_1\in L^2(\widetilde{Q})$. Let $\varphi\in C_c^\infty(\widetilde{Q})$. We compute

$$\begin{aligned}\int_{\widetilde{Q}}\widetilde{v}_1\partial_3\varphi\,\mathrm{d}x &= \int_{Q^{(1)}}\widetilde{v}_1\partial_3\varphi\,\mathrm{d}x+\int_{Q}\widetilde{v}_1\partial_3\varphi\,\mathrm{d}x+\int_{Q^{(2)}}\widetilde{v}_1\partial_3\varphi\,\mathrm{d}x\\ &= -\int_{Q^{(1)}}-\varphi(x)(\partial_2v_1)(x_1,x_2,2a_3^--x_3)\,\mathrm{d}x+\big[v_1(x_1,x_2,2a_3^--x_3)\varphi(x)\big]_{\Gamma_3^-}\\ &\quad-\int_{Q}\varphi(x)(\partial_3v_1)(x)\,\mathrm{d}x-\big[v_1(x)\varphi(x)\big]_{\Gamma_3^-}+\big[v_1(x)\varphi(x)\big]_{\Gamma_3^+}\\ &\quad-\int_{Q^{(2)}}-\varphi(x)\partial_3v_1)(x_1,x_2,2a_3^+-x_3)\,\mathrm{d}x-\big[v_1(x_1,x_2,2a_3^+-x_3)\varphi(x)\big]_{\Gamma_3^+}\\ &= -\int_{Q^{(1)}}-\varphi(x)(\partial_3v_1)(x_1,x_2,2a_3^--x_3)\,\mathrm{d}x-\int_{Q}\varphi(x)(\partial_3v_1)(x_1,x_2,x_3)\,\mathrm{d}x\end{aligned}$$

$$-\int_{Q^{(2)}} -\varphi(x)(\partial_3 v_1)(x_1,x_2,2a_3^+-x_3)\,\mathrm{d}x.$$

This shows that

$$(\partial_3\widetilde{v}_1)(x_1,x_2,x_3)=\begin{cases}-(\partial_3 v_1)(x_1,x_2,2a_3^- - x_3), & x_3\in(2a_3^- - a_3^+, a_3^-),\\ (\partial_3 v_1)(x_1,x_2,x_3), & x_3\in[a_3^- a_3^+],\\ -(\partial_3 v_1)(x_1,x_2,2a_3^+ - x_3), & x_3\in(a_3^+, 2a_3^+ - a_3^-),\end{cases}$$

belongs to $L^2(\widetilde{Q})$. Moreover, using $\partial_3 v_1 = 0$ on Γ_3 we get

$$\begin{aligned}\int_{\widetilde{Q}}(\partial_3\widetilde{v}_1)\partial_3\varphi\,\mathrm{d}x &= \int_{Q^{(1)}}(\partial_3\widetilde{v}_1)\partial_3\varphi\,\mathrm{d}x+\int_Q(\partial_3\widetilde{v}_1)\partial_3\varphi\,\mathrm{d}x+\int_{Q^{(2)}}(\partial_3\widetilde{v}_1)\partial_3\varphi\,\mathrm{d}x\\ &= -\int_{Q^{(1)}}\varphi(x)(\partial_{33}v_1)(x_1,x_2,2a_3^- - x_3)\,\mathrm{d}x\\ &\quad+\big[-(\partial_3 v_1)(x_1,x_2,2a_3^- - x_3)\varphi(x)\big]_{\Gamma_3^-}\\ &\quad-\int_Q\varphi(x)(\partial_{33}v_1)(x)\,\mathrm{d}x-\big[(\partial_3 v_1)(x)\varphi(x)\big]_{\Gamma_3^-}+\big[(\partial_3 v_1)(x)\varphi(x)\big]_{\Gamma_3^+}\\ &\quad-\int_{Q^{(2)}}\varphi(x)(\partial_{33}v_1)(x_1,x_2,2a_3^+ - x_3)\,\mathrm{d}x\\ &\quad-\big[-(\partial_3 v_1)(x_1,x_2,2a_3^+ - x_3)\varphi(x)\big]_{\Gamma_3^+}\\ &= -\int_{Q^{(1)}}\varphi(x)(\partial_{33}v_1)(x_1,x_2,2a_3^- - x_3)\,\mathrm{d}x-\int_Q\varphi(x)(\partial_{33}v_1)(x_1,x_2,x_3)\,\mathrm{d}x\\ &\quad-\int_{Q^{(2)}}\varphi(x)(\partial_{33}v_1)(x_1,x_2,2a_3^+ - x_3)\,\mathrm{d}x,\end{aligned}$$

so that

$$(\partial_{33}\widetilde{v}_1)(x_1,x_2,x_3)=\begin{cases}-(\partial_{33} v_1)(x_1,x_2,2a_3^- - x_3), & x\in(2a_3^- - a_3^+, a_3^-),\\ (\partial_{33} v_1)(x_1,x_2,x_3), & x_3\in[a_3^- a_3^+],\\ -(\partial_{33} v_1)(x_1,x_2.2a_2^+ - x_3), & x_3\in(a_3^+, 2a_3^+ - a_3^-),\end{cases}$$

is contained in $L^2(\widetilde{Q})$.

Furthermore, we have $\widetilde{v}_1 = 0$ on Γ_1, $\partial_2\widetilde{v}_1 = 0$ on Γ_2 and $\partial_3\widetilde{v}_1 = 0$ on Γ_3 due to the properties of Z. For $n \geq n_0$ we extend $\widetilde{v}_1$ by 0 to $\mathbb{R}^3$ and set

$$v_1^n(x_1,x_2,x_3) := \big(\rho_n^{(3)} * \widetilde{v}_1\big)\big|_{\widetilde{Q}}(x_1,x_2,x_3) = \int_{-1/n}^{1/n}\rho_n^{(3)}(t)\widetilde{v}_1(x_1,x_2,x_3-t)\,\mathrm{d}t\Big|_{\widetilde{Q}}$$

on $\widetilde{Q}$. Then v_1^n and $\partial_3 v_1^n$ belong to $H^2(\widetilde{Q})$. From the properties of $\widetilde{v}_1$ we derive in the same way as above for the traces of $\widetilde{u}_1^n$ that $v_1^n = 0$ on Γ_1 and $\partial_2 v_1^n = 0$ on Γ_2. Furthermore,

$$(\partial_3 v_1^n)(x_1,x_2,x_3)=\int_{-1/n}^{1/n}\rho_n^{(3)}(t)(\partial_3\widetilde{v}_1)(x_1,x_2,x_3-t)\,\mathrm{d}t,$$

so that

$$\begin{aligned}
&(\partial_3 v_1^n)(x_1, x_2, a_3^-) \\
&\quad = \int_{-1/n}^{0} \rho_n^{(3)}(t)(\partial_3 v_1)(x_1, x_2, a_3^- - t)\,\mathrm{d}t - \int_0^{1/n} \rho_n^{(3)}(t)(\partial_3 v_1)(x_1, x_2, a_3^- + t)\,\mathrm{d}t \\
&\quad = \int_0^{1/n} \rho_n^{(3)}(-s)(\partial_3 v_1)(x_1, x_2, a_3^- + s)\,\mathrm{d}s - \int_0^{1/n} \rho_n^{(3)}(t)(\partial_3 v_1)(x_1, x_2, a_3^- + t)\,\mathrm{d}t \\
&\quad = 0,
\end{aligned}$$

due to the symmetry of $\rho_n^{(3)}$. Analogously, we get $\partial_3 v_1^n(x_1, x_2, a_3^+) = 0$, so that together $\partial_3 v_1 = 0$ on Γ_3. We have $v_1^n \to \widetilde{v}_1$ in $H^2(\widetilde{Q})$ as $n \to \infty$ and therefore $v_1^n|_Q \to v_1$ in $H^2(Q)$ as $n \to \infty$. □

Under the assumption $\varepsilon, \mu, \sigma \in W^{2,3}(Q)$ we set

$$\begin{aligned}
\kappa_Z := {} & \frac{7\,\|\nabla\sigma\|_{L^\infty}}{4\delta} + \frac{7\,\|\sigma\|_{L^\infty}\,\|\nabla\varepsilon\|_{L^\infty}}{4\delta^2} + \frac{6\,\|\nabla\varepsilon\|_{L^\infty} + 6\,\|\nabla\mu\|_{L^\infty}}{2\delta^2} \\
& + \frac{9C_{H^1\hookrightarrow L^6}\,\|\sigma\|_{W^{2,3}}}{4\delta} + \frac{9C_{H^1\hookrightarrow L^6}\,\|\sigma\|_{L^\infty}\,\|\varepsilon\|_{W^{2,3}}}{4\delta^2} \\
& + \frac{9\,\|\nabla\sigma\|_{L^\infty}\,\|\nabla\varepsilon\|_{L^\infty}}{2\delta^2} + \frac{9\,\|\sigma\|_{L^\infty}\,\|\nabla\varepsilon\|_{L^\infty}^2}{2\delta^3} \\
& + \frac{5C_{H^1\hookrightarrow L^6}\,\|\varepsilon\|_{W^{2,3}} + 5C_{H^1\hookrightarrow L^6}\,\|\mu\|_{W^{2,3}}}{\delta^2} + \frac{9\,\|\nabla\varepsilon\|_{L^\infty}^2 + 9\,\|\nabla\mu\|_{L^\infty}^2}{\delta^3},
\end{aligned}$$

where $C_{H^1\hookrightarrow L^6}$ denotes the Sobolev embedding constant from $H^1(Q)$ to $L^6(Q)$.

Lemma 8.10. *Let $\varepsilon, \mu, \sigma \in W^{2,3}(Q)$. Then the operators $A_Z - \kappa_Z I$ and $B_Z - \kappa_Z I$ are dissipative on Z.*

Proof:
Let $(u, v) \in D(A_Z)$. With integration by parts we see

$$\begin{aligned}
&\sum_{j,k=1}^{3} \int_Q \big(\partial_{jk} C_1 v \cdot \partial_{jk} u + \partial_{jk} C_2 u \cdot \partial_{jk} v\big)\,\mathrm{d}x \\
&\quad = \sum_{j,k=1}^{3} \int_Q \big(\partial_{jk2} v_3\,\partial_{jk} u_1 + \partial_{jk3} v_1\,\partial_{jk} u_2 + \partial_{jk1} v_2\,\partial_{jk} u_3 \\
&\qquad + \partial_{jk3} u_2\,\partial_{jk} v_1 + \partial_{jk1} u_3\,\partial_{jk} v_2 + \partial_{jk2} u_1\,\partial_{jk} v_3\big)\,\mathrm{d}x \\
&\quad = 0,
\end{aligned}$$

where we have used the boundary properties of from Corollary 8.8 and of the definition of $D(A_Z)$ to get rid of the boundary integrals. As in the proof of Lemma 7.2 we have for example

$$\|(\partial_{jk}\sigma)u\|_{L^2} \le C_{H^1\hookrightarrow L^6}\,\|\sigma\|_{W^{2,3}}\,\|u\|_{H^1}$$

for all $(u,v) \in D(A_Z)$. Together with (7.26), (7.28), (8.2) and Young's inequality we thus estimate

$$
\begin{aligned}
&\operatorname{Re}\big(A(u,v) \mid (u,v)\big)_Z \\
&\quad= \int_Q \Big(-\frac{\sigma\varepsilon}{2\varepsilon} |u|^2 + \frac{\varepsilon}{\varepsilon} C_1 v \cdot u + \frac{\mu}{\mu} C_2 u \cdot v \\
&\qquad - \varepsilon \sum_{j=1}^3 \partial_j \big(\tfrac{\sigma}{2\varepsilon} u\big) \cdot \partial_j u + \varepsilon \sum_{j=1}^3 \partial_j \big(\tfrac{1}{\varepsilon} C_1 v\big) \cdot \partial_j u + \mu \sum_{j=1}^3 \partial_j \big(\tfrac{1}{\mu} C_2 u\big) \cdot \partial_j v \\
&\qquad - \varepsilon \sum_{j,k=1}^3 \partial_{jk} \big(\tfrac{\sigma}{2\varepsilon} u\big) \cdot \partial_{jk} u + \varepsilon \sum_{j,k=1}^3 \partial_{jk} \big(\tfrac{1}{\varepsilon} C_1 v\big) \cdot \partial_{jk} u + \mu \sum_{j,k=1}^3 \partial_{jk} \big(\tfrac{1}{\mu} C_2 u\big) \cdot \partial_{jk} v \Big) \,\mathrm{d}x \\
&\quad= \int_Q -\frac{\sigma}{2} |u|^2 \,\mathrm{d}x - \int_Q \frac{\sigma}{2} |\partial u|^2 \,\mathrm{d}x - \sum_{j=1}^3 \int_Q \big(\tfrac{\partial_j \sigma}{2} - \tfrac{\sigma \partial_j \varepsilon}{2\varepsilon}\big) u \cdot \partial_j u \,\mathrm{d}x \\
&\qquad - \sum_{j=1}^3 \int_Q \frac{\partial_j \varepsilon}{\varepsilon} C_1 v \cdot \partial_j u \,\mathrm{d}x - \sum_{j=1}^3 \int_Q \frac{\partial_j \mu}{\mu} C_2 v \cdot \partial_j u \,\mathrm{d}x \\
&\qquad - \frac{\sigma}{2} \sum_{j,k=1}^3 \int_Q |\partial_{jk} u|^2 \,\mathrm{d}x - \sum_{j,k=1}^3 \int_Q \Big(\big(\tfrac{\partial_j \sigma}{2} - \tfrac{\sigma \partial_j \varepsilon}{2\varepsilon}\big) \partial_k u + \big(\tfrac{\partial_k \sigma}{2} - \tfrac{\sigma \partial_k \varepsilon}{2\varepsilon}\big) \partial_j u \Big) \cdot \partial_{jk} u \,\mathrm{d}x \\
&\qquad - \sum_{j,k=1}^3 \int_Q \big(\tfrac{\partial_{jk} \sigma}{2} - \tfrac{\partial_j \sigma \partial_k \varepsilon}{2\varepsilon} - \tfrac{\partial_k \sigma \partial_j \varepsilon}{2\varepsilon} - \tfrac{\sigma \partial_{jk} \varepsilon}{2\varepsilon} + \tfrac{\sigma (\partial_j \varepsilon) \partial_k \varepsilon}{\varepsilon^2}\big) u \cdot \partial_{jk} u \,\mathrm{d}x \\
&\qquad + \sum_{j,k=1}^3 \int_Q \big(-\tfrac{\partial_{jk} \varepsilon}{\varepsilon} + \tfrac{2(\partial_j \varepsilon) \partial_k \varepsilon}{\varepsilon^2}\big) C_1 v \cdot \partial_{jk} u \,\mathrm{d}x + \sum_{j,k=1}^3 \int_Q \big(-\tfrac{\partial_j \varepsilon}{\varepsilon} \partial_k C_2 v - \tfrac{\partial_k \varepsilon}{\varepsilon} \partial_j C_1 v\big) \cdot \partial_{jk} u \,\mathrm{d}x \\
&\qquad + \sum_{j,k=1}^3 \int_Q \big(-\tfrac{\partial_{jk} \mu}{\mu} + \tfrac{2(\partial_j \mu) \partial_k \mu}{\mu^2}\big) C_2 u \cdot \partial_{jk} v \,\mathrm{d}x \\
&\qquad + \sum_{j,k=1}^3 \int_Q \big(-\tfrac{\partial_j \mu}{\mu} \partial_k C_2 u - \tfrac{\partial_k \mu}{\mu} \partial_j C_2 u\big) \cdot \partial_{jk} v \,\mathrm{d}x \\
&\quad\leq \Big(\frac{\|\nabla \sigma\|_{L^\infty}}{4\delta} + \frac{\|\sigma\|_{L^\infty} \|\nabla \varepsilon\|_{L^\infty}}{4\delta^2} \Big) \int_Q \big(3\varepsilon |u|^2 + \varepsilon |\partial u|^2\big) \,\mathrm{d}x \\
&\qquad + \frac{\|\nabla \varepsilon\|_{L^\infty} + \|\nabla \mu\|_{L^\infty}}{2\delta^2} \int_Q \big(3\varepsilon |\partial u|^2 + 3\mu |\partial v|^2\big) \,\mathrm{d}x \\
&\qquad + \Big(\frac{\|\nabla \sigma\|_{L^\infty}}{2\delta} + \frac{\|\sigma\|_{L^\infty} \|\nabla \varepsilon\|_{L^\infty}}{2\delta^2} \Big) \int_Q \big(3\varepsilon |\partial u|^2 + \varepsilon \big|D^2 u\big|^2\big) \,\mathrm{d}x \\
&\qquad + C_{H^1 \hookrightarrow L^6} \Big(\frac{\|\sigma\|_{W^{2,3}}}{4\delta} + \frac{\|\sigma\|_{L^\infty} \|\varepsilon\|_{W^{2,3}}}{4\delta^2} \Big) \int_Q \big(9\varepsilon |u|^2 + 9\varepsilon |\partial u|^2 + \varepsilon \big|D^2 u\big|^2\big) \,\mathrm{d}x \\
&\qquad + \Big(\frac{\|\nabla \sigma\|_{L^\infty} \|\nabla \varepsilon\|_{L^\infty}}{2\delta^2} + \frac{\|\sigma\|_{L^\infty} \|\nabla \varepsilon\|_{L^\infty}^2}{2\delta^3} \Big) \int_Q \big(9\varepsilon |u|^2 + \varepsilon \big|D^2 u\big|^2\big) \,\mathrm{d}x
\end{aligned}
$$

$$
\begin{aligned}
&+ C_{H^1\hookrightarrow L^6}\frac{\|\varepsilon\|_{W^{2,3}}+\|\mu\|_{W^{2,3}}}{2\delta^2}\int_Q \left(9\varepsilon\left|\partial u\right|^2+9\mu\left|\partial v\right|^2+10\varepsilon\left|D^2u\right|^2+10\mu\left|D^2v\right|^2\right)\mathrm{d}x\\
&+ \frac{\|\nabla\varepsilon\|_{L^\infty}^2+\|\nabla\mu\|_{L^\infty}^2}{\delta^3}\int_Q \left(9\varepsilon\left|\partial u\right|^2+9\mu\left|\partial v\right|^2+\varepsilon\left|D^2u\right|^2+\mu\left|D^2v\right|^2\right)\mathrm{d}x\\
&+ \frac{\|\nabla\varepsilon\|_{L^\infty}+\|\nabla\mu\|_{L^\infty}}{\delta^2}\int_Q \left(3\varepsilon\left|D^2u\right|^2+3\mu\left|D^2v\right|^2\right)\mathrm{d}x\\
&\leq \kappa_Z\left\|(u,v)\right\|_Z^2,
\end{aligned}
$$

where the norm of the Jacobian matrix and the matrix of the second derivatives is the Frobenius norm.

Let $(u,v)\in D(B_Z)$. In the same way as for A_Z we see

$$
\sum_{j,k=1}^{3}\int_Q \left(\partial_{jk}C_2v\cdot\partial_{jk}u+\partial_{jk}C_1u\cdot\partial_{jk}v\right)\mathrm{d}x=0
$$

and estimate

$$
\begin{aligned}
&\mathrm{Re}\big(B(u,v)\mid(u,v)\big)_Z\\
&\quad=\int_Q\Big(-\frac{\sigma\varepsilon}{2\varepsilon}\left|u\right|^2+\frac{\varepsilon}{\varepsilon}C_2v\cdot u+\frac{\mu}{\mu}C_1u\cdot v\\
&\qquad-\varepsilon\sum_{j=1}^{3}\partial_j\big(\tfrac{\sigma}{2\varepsilon}u\big)\cdot\partial_ju+\varepsilon\sum_{j=1}^{3}\partial_j\big(\tfrac{1}{\varepsilon}C_2v\big)\cdot\partial_ju+\mu\sum_{j=1}^{3}\partial_j\big(\tfrac{1}{\mu}C_1u\big)\cdot\partial_jv\\
&\qquad-\varepsilon\sum_{j,k=1}^{3}\partial_{jk}\big(\tfrac{\sigma}{2\varepsilon}u\big)\cdot\partial_{jk}u+\varepsilon\sum_{j,k=1}^{3}\partial_{jk}\big(\tfrac{1}{\varepsilon}C_2v\big)\cdot\partial_{jk}u+\mu\sum_{j,k=1}^{3}\partial_{jk}\big(\tfrac{1}{\mu}C_1u\big)\cdot\partial_{jk}v\Big)\,\mathrm{d}x\\
&\quad\leq\kappa_Z\left\|(u,v)\right\|_Z^2,
\end{aligned}
$$

which finishes the proof. □

Lemma 8.11. *Let $\varepsilon,\sigma\in W^{2,3}(Q)$, $\mu\in C^2(\overline{Q})$ and $\partial_\nu\varepsilon=\partial_\nu\mu=\partial_\nu\sigma=0$ on Γ. Then the operators $(1+\kappa_Z)I-A_Z$ and $(1+\kappa_Z)I-B_Z$ have a dense range in Z.*

Proof:
We first deal with the operator $(1+\kappa_Z)I-A_Z$. Having the denseness of $D(A_Z)$ in Z by Lemma 8.9 in mind, let $(f,g)\in D(A_Z)\subseteq D(A_Y)$. As in the proof of Lemma 8.6 we have equation (8.4) with κ_Z instead of κ_Y. Due to Lemma 8.6 there exists fields $(u,v)\in D(A_Y)$ that solves (8.4) and thus satisfies in particular

$$
\big((1+\kappa_Z)\varepsilon+\tfrac{\sigma}{2}\big)u_1-\partial_2v_3=\varepsilon f_1, \tag{8.6a}
$$

$$
\mu(1+\kappa_Z)v_3-\partial_2u_1=\mu g_3. \tag{8.6b}
$$

From the definition of L from the proof of Lemma 8.6 with κ_Y replaced by κ_Z we derive from these equations the identity

$$Lu_1 = \big((1+\kappa_Z)\varepsilon + \tfrac{\sigma}{2}\big)u_1 - \frac{1}{1+\kappa_Z}\partial_2\big(\tfrac{1}{\mu}\partial_2 u_1\big) = \varepsilon f_1 + \frac{1}{1+\kappa_Z}\partial_2 g_3 =: h_1. \tag{8.7}$$

Due to the properties of $D(A_Y)$ we know that u_1, v_3, $\partial_2 u_1$ and $\partial_2 v_3$ are contained in $H^1(Q)$, $u_1 = 0$ on $\Gamma_2 \cup \Gamma_3$, $v_3 = 0$ on Γ_3, $\partial_2 u_1 = 0$ on Γ_3 and $\partial_2 v_3 = 0$ on $\Gamma_2 \cup \Gamma_3$. The properties of $D(A_Z)$ furthermore give that f_1, g_3, $\partial_2 f_1$ and $\partial_2 g_3$ belong to $H^2(Q)$, $f_1 = 0$ on $\Gamma_2 \cup \Gamma_3$, $g_3 = 0$ on Γ_3, $\partial_1 f_1 = 0$ on Γ_1, $\partial_{22} f_1 = 0$ on Γ_2 and $\partial_j g_3 = 0$ on Γ_j for $j \in \{1,2\}$. So, h_1 is contained in $H^2(Q)$ and $h_1 = 0$ on Γ_2, due to Lemma 7.5.

From $\partial_2 u_1 \in H^1(Q)$ and $\mu \in W^{1,\infty}(Q)$ we infer $D_2\partial_{jk}u_1 \in H^{-2}(Q)$ for all $j,k \in \{1,2,3\}$. Let $\varphi \in H_0^3(Q)$ and $j,k \in \{1,2,3\}$. Using the regularity of the coefficients, we can thus estimate

$$\begin{aligned}
\langle L\partial_{jk}u_1, \varphi\rangle_{H^{-2}\times H_0^2} &= \big\langle \partial_{jk}u_1, \big((1+\kappa_Z)\varepsilon + \tfrac{\sigma}{2}\big)\varphi\big\rangle_{H^{-1}\times H_0^1} - \frac{1}{1+\kappa_Z}\Big\langle \partial_2 \tfrac{1}{\mu}\partial_2\partial_{jk}u_1, \varphi\Big\rangle_{H^{-2}\times H_0^2} \\
&= \int_Q u_1\Big(\big(\partial_{jk}\big((1+\kappa_Z)\varepsilon + \tfrac{\sigma}{2}\big)\big)\varphi + \big(\partial_j\big((1+\kappa_Z)\varepsilon + \tfrac{\sigma}{2}\big)\big)\partial_k\varphi \\
&\qquad + \big(\partial_k\big((1+\kappa_Z)\varepsilon + \tfrac{\sigma}{2}\big)\big)\partial_j\varphi + \big((1+\kappa_Z)\varepsilon + \tfrac{\sigma}{2}\big)\partial_{jk}\varphi\Big)\,\mathrm{d}x \\
&\quad + \int_Q (\partial_2 u_1)\Big(\big(\partial_{jk}\tfrac{1}{\mu}\big)\partial_2\varphi + \big(\partial_j\tfrac{1}{\mu}\big)\partial_k\partial_2\varphi + \big(\partial_k\tfrac{1}{\mu}\big)\partial_j\partial_2\varphi + \tfrac{1}{\mu}\partial_2\partial_{jk}\varphi\Big)\,\mathrm{d}x \\
&= \int_Q Lu_1\partial_{jk}\varphi\,\mathrm{d}x + \int_Q u_1\big(\partial_{jk}\big((1+\kappa_Z)\varepsilon + \tfrac{\sigma}{2}\big)\big)\varphi\,\mathrm{d}x \\
&\quad + \int_Q u_1\Big(\big(\partial_j\big((1+\kappa_Z)\varepsilon + \tfrac{\sigma}{2}\big)\big)\partial_k\varphi + \big(\partial_k\big((1+\kappa_Z)\varepsilon + \tfrac{\sigma}{2}\big)\big)\partial_j\varphi\Big)\,\mathrm{d}x \\
&\quad + \int_Q (\partial_2 u_1)\Big(\big(\partial_j\tfrac{1}{\mu}\big)\partial_k\partial_2\varphi + \big(\partial_k\tfrac{1}{\mu}\big)\partial_j\partial_2\varphi + \big(\partial_{jk}\tfrac{1}{\mu}\big)\partial_2\varphi\Big)\,\mathrm{d}x \\
&= \int_Q (\partial_{jk}h_1)\varphi\,\mathrm{d}x - \int_Q u_1\big(\partial_{jk}\big((1+\kappa_Z)\varepsilon + \tfrac{\sigma}{2}\big)\big)\varphi\,\mathrm{d}x \\
&\quad - \int_Q \Big((\partial_k u_1)\big(\partial_j\big((1+\kappa_Z)\varepsilon + \tfrac{\sigma}{2}\big)\big) + (\partial_j u_1)\big(\partial_k\big((1+\kappa_Z)\varepsilon + \tfrac{\sigma}{2}\big)\big)\Big)\varphi\,\mathrm{d}x \\
&\quad + \Big\langle \partial_2\partial_k\big((\partial_2 u_1)\big(\partial_j\tfrac{1}{\mu}\big)\big) + \partial_2\partial_j\big((\partial_2 u_1)\big(\partial_k\tfrac{1}{\mu}\big)\big), \varphi\Big\rangle_{D(\partial_2)^*\times D(\partial_2)} \\
&\quad - \Big\langle \partial_2\big((\partial_2 u_1)\big(\partial_{jk}\tfrac{1}{\mu}\big)\big), \varphi\Big\rangle_{D(\partial_2)^*\times D(\partial_2)}.
\end{aligned}$$

As in the proof of Lemma 8.6, by the density of $H_0^3(Q)$ in $D(\partial_2)$ this identity holds true for all $\varphi \in D(\partial_2)$ and so

$$\begin{aligned}
L\partial_{jk}u_1 &= \partial_{jk}h_1 - \big((1+\kappa_Z)\partial_{jk}\varepsilon + \tfrac{\partial_{jk}\sigma}{2}\big)u_1 - \big((1+\kappa_Z)\partial_j\varepsilon + \tfrac{\partial_j\sigma}{2}\big)\partial_k u_1 \\
&\quad - \big((1+\kappa_Z)\partial_k\varepsilon + \tfrac{\partial_k\sigma}{2}\big)\partial_j u_1 + \partial_2\big(\big(\partial_{jk}\tfrac{1}{\mu}\big)\partial_2 u_1\big) \\
&\quad + \partial_2\big(\big(\partial_j\tfrac{1}{\mu}\big)\partial_{2k}u_1\big) + \partial_2\big(\big(\partial_k\tfrac{1}{\mu}\big)\partial_{2j}u_1\big)
\end{aligned}$$

$$=: \psi_1(h_1) \in D(\partial_2)^*$$

first in $H^{-2}(Q)$, and then that it holds true even in $D(\partial_2)^*$ since all summands on the right-hand side are in $D(\partial_2)^*$.
As in the proof of Lemma 8.6, we now conclude $\partial_{jk}u_1 = L^{-1}\psi_1(h_1) \in D(\partial_2)$. Since $j,k \in \{1,2,3\}$ were arbitrary and the weak derivatives of first order can be treated as in the proof of Lemma 8.6, we thus have that u_1 and $\partial_2 u_1$ are contained in $H^2(Q)$. With (8.6) and Lemma 7.2 this gives that $\partial_2 v_3$ and v_3 belong to $H^2(Q)$. From (8.7), Lemma 7.5 and $h_1 = 0$ on Γ_2 we infer

$$\partial_2\big(\tfrac{1}{\mu}\partial_2 u_1\big) = (1+\kappa_Z)\big(\varepsilon(1+\kappa_Z) + \tfrac{\sigma}{2}\big)u_1 - (1+\kappa_Z)h_1 = 0$$

on Γ_2, so that using $\partial_\nu \mu = 0$ on Γ we have $\partial_{22}u_1 = 0$ on Γ_2. It remains to prove that $\partial_1 v_3 = 0$ and $\partial_1 u_1 = 0$ on Γ_1. Lemma 7.10 and the identity $\partial_1 g_3 = 0$ on Γ_1 imply $\partial_2\partial_1 g_3 = 0$ on Γ_1. With $\partial_\nu \varepsilon = 0$ on Γ_1 and $\partial_1 f_1 = 0$ on Γ_1 we thus deduce

$$\partial_1 h_1 = (\partial_1\varepsilon)f_1 + \varepsilon\partial_1 f_1 + \partial_1\partial_2 g_3 = 0$$

on Γ_1. moreover, the conditions $\partial_\nu\varepsilon = \partial_\nu\sigma = 0$ on Γ yield

$$\partial_1\big((1+\kappa_Z)\varepsilon + \tfrac{\sigma}{2}\big) = 0$$

on Γ_1. Since $\mu \in C^2(\overline{Q})$ and $\partial_1\mu = 0$ on Γ_1 we have that $\partial_{21}\mu = 0$ on Γ_1 and thus

$$\partial_{12}\frac{1}{\mu} = -\frac{\partial_{12}\mu}{\mu^2} + \frac{2(\partial_1\mu)(\partial_2\mu)}{\mu^3} = 0$$

on Γ_1. Hence,

$$\partial_2\Big(\big(\partial_1\tfrac{1}{\mu}\big)\partial_2 u_1\Big) = \big(\partial_{12}\tfrac{1}{\mu}\big)\partial_2 u_1 - \frac{\partial_1\mu}{\mu}\partial_{22}u_1 = 0$$

on Γ_1, using that $\partial_2 u_1$ belongs to $H^2(Q)$. Taking the last above facts into account, we deduce from an analogon of equation (8.5) that the function

$$L\partial_1 u_1 = \big(\partial_1\big((1+\kappa_Z)\varepsilon - \tfrac{\sigma}{2}\big)\big)u_1 + \frac{1}{1+\kappa_Z}\partial_2\big(\big(\partial_1\tfrac{1}{\mu}\big)\partial_2 u_1\big) =: \varphi_1 \in H^1(Q)$$

vanishes on Γ_1. Set $\varphi_1^n := \chi_n^{(1)}\varphi_1 \in H^1(Q)$. We have $\varphi_1^n = 0$ on Γ_1 and $\varphi_1^n \to \varphi_1$ in $L^2(Q)$ as $n \to \infty$. As in the proof of Lemma 8.6 we now infer $L^{-1}\varphi_1^n = 0$ on Γ_1 (even on a neighbourhood of Γ_1) and by the continuity of L^{-1} on $L^2(Q)$ that $L^{-1}\varphi_1^n \to L^{-1}\varphi_1 = \partial_1 u_1$ in $H^1(Q)$ as $n \to \infty$. Thus, $\partial_1 u_1 = 0$ on Γ_1. From this we conclude with Lemma 7.10 that $\partial_{12}u_1 = 0$ on Γ_1 and hence with (8.6b), divided by μ, $\partial_\nu\mu = 0$ on Γ and Lemma 7.5 that $\partial_1 v_3 = 0$ on Γ_1. Treating the components u_2, u_3, v_1, v_2 as u_1 and v_3, respectively, we have altogether $(u,v) \in D(A_Z)$.

Replacing (8.6) by

$$\begin{aligned}\big((1+\kappa_Z)\varepsilon+\tfrac{\sigma}{2}\big)u_1+\partial_3 v_2&=\varepsilon f_1,\\ \mu(1+\kappa_Z)v_2+\partial_3 u_1&=\mu g_2.\end{aligned}$$

and (8.7) by

$$Lu_1=\big((1+\kappa_Z)\varepsilon+\tfrac{\sigma}{2}\big)u_1-\frac{1}{1+\kappa_Z}\partial_3\big(\tfrac{1}{\mu}\partial_3 u_1\big)=\varepsilon f_1+\partial_3 g_2=:h_1,$$

the statement involving the operator B_Z is shown in the same way. □

With the same proof as for Proposition 8.7, invoking Lemma 8.9, 8.10 and 8.11, one sees the following proposition on the resolvents of A_Z and B_Z.

Proposition 8.12. *Let $\varepsilon,\sigma\in W^{2,3}(Q)$, $\mu\in C^2(\overline{Q})$ and $\partial_\nu\varepsilon=\partial_\nu\mu=\partial_\nu\sigma=0$ on Γ.*

(a) The operators A_Z and B_Z generate C_0-semigroups on Z whose norms are bounded by $e^{\kappa_Z t}$. The restrictions of $(I-\tau A_Y)^{-1}$ and $(I-\tau B_Y)^{-1}$ to Z are the operators $(I-\tau A_Z)^{-1}$ and $(I-\tau B_Z)^{-1}$, respectively. The semigroup estimate implies

$$\big\|(I-\tau A_Z)^{-1}\big\|_{\mathcal{B}(Z)}\le\frac{1}{1-\tau\kappa_Z}\qquad\text{and}\qquad\big\|(I-\tau B_Z)^{-1}\big\|_{\mathcal{B}(Z)}\le\frac{1}{1-\tau\kappa_Z}$$

for all $0<\tau<\frac{1}{\kappa_Z}$, which means in particular

$$\big\|(I-\tau A_Z)^{-1}\big\|_{\mathcal{B}(Z)}\le 2\qquad\text{and}\qquad\big\|(I-\tau B_Z)^{-1}\big\|_{\mathcal{B}(Z)}\le 2$$

for all $0<\tau\le\frac{1}{2\kappa_Z}$. Moreover, the operators $A_Z-\kappa_Z I$ and $B_Z-\kappa_Z I$ are maximal dissipative on Z.

(b) We define the function

$$\gamma_\tau(z):=\frac{1+\tau z}{1-\tau z}$$

on $\mathbb{C}\setminus\{\frac{1}{\tau}\}$. Then there exists a $\widetilde{\tau}\in\big(0,\frac{1}{\kappa_Z}\big)$ such that

$$\begin{aligned}\|\gamma_\tau(A_Z)\|_{\mathcal{B}(Z)}&\le e^{3\kappa_Z\tau},&\qquad\|\gamma_\tau(B_Z)\|_{\mathcal{B}(Z)}&\le e^{3\kappa_Z\tau},\\ \|\gamma_\tau(A_Z^*)\|_{\mathcal{B}(Z)}&\le e^{3\kappa_Z\tau},&\qquad\|\gamma_\tau(B_Z^*)\|_{\mathcal{B}(Z)}&\le e^{3\kappa_Y\tau}\end{aligned}$$

for all $0<\tau<\widetilde{\tau}$.

8.4. The ADI splitting scheme

Let $\tau > 0$. We set $t_n := n\tau$ for $n \in \mathbb{N}_0$ and assume $(\mathbf{J}_0(t), 0) \in D(A)$ for all $t \geq 0$. The *alternating direction implicit (ADI) splitting scheme* $S^I_{\tau,n+1}$ we investigate is given by

$$\begin{aligned} S^I_{\tau,n+1} w := (I - \tfrac{\tau}{2}B)^{-1}(I + \tfrac{\tau}{2}A)\cdot \\ \cdot \Big[(I - \tfrac{\tau}{2}A)^{-1}(I + \tfrac{\tau}{2}B)w - \tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1}), 0)\Big] \end{aligned} \tag{8.8}$$

for $w \in D(B)$, as introduced in Section 7.1. Proposition 8.1 shows that the resolvents in (8.8) exist. Thus, the splitting scheme is well-defined.

We divide the splitting scheme (8.8) into the two parts

$$S^{I,(1)}_\tau w_1 := (I - \tfrac{\tau}{2}A)^{-1}(I + \tfrac{\tau}{2}B)w_1 \in D(A) \quad \text{for } w_1 \in D(B) \quad \text{and} \tag{8.9a}$$

$$S^{I,(2)}_\tau w_2 := (I - \tfrac{\tau}{2}B)^{-1}(I + \tfrac{\tau}{2}A)w_2 \in D(B) \quad \text{for } w_2 \in D(A), \tag{8.9b}$$

which together give

$$S^I_{\tau,n+1} w = S^{I,(2)}_\tau \Big[S^{I,(1)}_\tau w - \tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1}), 0)\Big] \in D(B) \tag{8.9c}$$

for $w \in D(B)$. Both $S^{I,(1)}_\tau$ and $S^{I,(2)}_\tau$ contain an implicit part that results in a linear system with three coupled equations.

For a better overview concerning the physical meaning of the variables we switch our notation to variables containing the electric and the magnetic field. For $n \in \mathbb{N}_0$ and $(\mathbf{E}_0, \mathbf{H}_0) \in D(B)$ this gives

$$(\mathbf{E}_{n+1/2}, \mathbf{H}_{n+1/2}) := S^{I,(1)}_\tau (\mathbf{E}_n, \mathbf{H}_n) \in D(A), \tag{8.10a}$$

$$(\mathbf{E}_{n+1}, \mathbf{H}_{n+1}) := S^{I,(2)}_\tau \Big[(\mathbf{E}_{n+1/2}, \mathbf{H}_{n+1/2}) - \tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1}), 0)\Big] \in D(B) \tag{8.10b}$$

and

$$(\mathbf{E}_n, \mathbf{H}_n) := S^I_{\tau,n} \cdots S^I_{\tau,1} (\mathbf{E}_0, \mathbf{H}_0). \tag{8.10c}$$

Taking Proposition 8.7 and 8.12 into account, we get the following statements in the H^1- and the H^2-setting.

Remark 8.13. *(a) If $(\boldsymbol{E}_0, \boldsymbol{H}_0) \in D(B_Y)$ and $(\boldsymbol{J}_0(t), 0) \in D(A_Y)$ for all $t \in \mathbb{R}$, then $(\boldsymbol{E}_n, \boldsymbol{H}_n) \in D(B_Y)$ and $(\boldsymbol{E}_{n+1/2}, \boldsymbol{H}_{n+1/2}) \in D(A_Y)$ for all $n \in \mathbb{N}_0$.*

(b) Let $\varepsilon, \sigma \in W^{2,3}(Q)$, $\mu \in C^2(\overline{Q})$ and $\partial_\nu \varepsilon = \partial_\nu \mu = \partial_\nu \sigma = 0$ on Γ. If $(\boldsymbol{E}_0, \boldsymbol{H}_0) \in D(B_Z)$ and $(\boldsymbol{J}_0(t), 0) \in D(A_Z)$ for all $t \in \mathbb{R}$, then $(\boldsymbol{E}_n, \boldsymbol{H}_n) \in D(B_Z)$ and $(\boldsymbol{E}_{n+1/2}, \boldsymbol{H}_{n+1/2}) \in D(A_Z)$ for all $n \in \mathbb{N}_0$.

8.5. The efficiency of the ADI splitting scheme

In the computation of a numerical solution to the Maxwell equations (7.1) one has to solve implicit systems of linear equations. They arise from the two resolvents in the ADI splitting scheme (8.8). As in [75] and [37] we replace (8.10) by equivalent formulations such that the linear systems of three-dimensional equations decouple into three one-dimensional equations each. So, they can be solved in an efficient way. This important property of the ADI scheme is the main advantage of the present method over most other implicit methods.

For $\lambda \in \{\varepsilon, \mu\}$ we define the operators

$$\begin{aligned} D_\lambda^{(1)} : \{u \in L^2(Q)^3 \mid C_2 u \in H^1(Q)^3, \\ u_1 = 0 \text{ on } \Gamma_2,\ u_2 = 0 \text{ on } \Gamma_3,\ u_3 = 0 \text{ on } \Gamma_1\} \to L^2(Q)^3 \end{aligned}$$

and

$$\begin{aligned} D_\lambda^{(2)} : \{u \in L^2(Q)^3 \mid C_1 u \in H^1(Q)^3, \\ u_1 = 0 \text{ on } \Gamma_3,\ u_2 = 0 \text{ on } \Gamma_1,\ u_3 = 0 \text{ on } \Gamma_2\} \to L^2(Q)^3 \end{aligned}$$

by

$$D_\lambda^{(1)} := C_1 \tfrac{1}{\lambda} C_2 = \begin{pmatrix} \partial_2 \frac{1}{\lambda} \partial_2 & 0 & 0 \\ 0 & \partial_3 \frac{1}{\lambda} \partial_3 & 0 \\ 0 & 0 & \partial_1 \frac{1}{\lambda} \partial_1 \end{pmatrix} \tag{8.11a}$$

and

$$D_\lambda^{(2)} := C_2 \tfrac{1}{\lambda} C_1 = \begin{pmatrix} \partial_3 \frac{1}{\lambda} \partial_3 & 0 & 0 \\ 0 & \partial_1 \frac{1}{\lambda} \partial_1 & 0 \\ 0 & 0 & \partial_2 \frac{1}{\lambda} \partial_2 \end{pmatrix}. \tag{8.11b}$$

Let $(\mathbf{J}_0(t), 0) \in D(A_Y)$ for all $t \geq 0$. Starting with $(\mathbf{E}_n, \mathbf{H}_n) \in D(B_Y)$ for an $n \in \mathbb{N}$ we have due to (8.9) and (8.10) in $H^1(Q)^3$ for $n \in \mathbb{N}_0$ that

$$\begin{aligned} \big(1 + \tfrac{\sigma\tau}{4\varepsilon}\big)\mathbf{E}_{n+1/2} &= \big(1 - \tfrac{\sigma\tau}{4\varepsilon}\big)\mathbf{E}_n - \tfrac{\tau}{2\varepsilon} C_2 \mathbf{H}_n + \tfrac{\tau}{2\varepsilon} C_1 \mathbf{H}_{n+1/2}, \\ \mathbf{H}_{n+1/2} &= \mathbf{H}_n - \tfrac{\tau}{2\mu} C_1 \mathbf{E}_n + \tfrac{\tau}{2\mu} C_2 \mathbf{E}_{n+1/2}, \end{aligned}$$

with $(\mathbf{E}_{n+1/2}, \mathbf{H}_{n+1/2}) \in D(A_Y)$. Plugging the second equation into the first one, we eliminate $\mathbf{H}_{n+1/2}$ therein and get

$$\big((1 + \tfrac{\sigma\tau}{4\varepsilon})I - \tfrac{\tau^2}{4\varepsilon} D_\mu^{(1)}\big)\mathbf{E}_{n+1/2} = \big(1 - \tfrac{\sigma\tau}{4\varepsilon}\big)\mathbf{E}_n + \tfrac{\tau}{2\varepsilon} \operatorname{curl} \mathbf{H}_n - \tfrac{\tau^2}{4\varepsilon} C_1 \tfrac{1}{\mu} C_1 \mathbf{E}_n,$$

$$\mathbf{H}_{n+1/2} = \mathbf{H}_n - \tfrac{\tau}{2\mu} C_1 \mathbf{E}_n + \tfrac{\tau}{2\mu} C_2 \mathbf{E}_{n+1/2}$$

in $L^2(Q)^3$. The representation of $D_\mu^{(1)}$ as diagonal matrix in (8.11) shows that the implicit part of the first equation decouples into three independent equations, so that the first half-step of the spitting scheme can be computed efficiently.

Again with (8.9) and (8.10) we see that for $(\mathbf{E}_{n+1}, \mathbf{H}_{n+1}) \in D(B_Y)$ and $n \in \mathbb{N}$ we have in $H^1(Q)^3$ that

$$\begin{aligned}
\big(1 + \tfrac{\sigma\tau}{4\varepsilon}\big)\mathbf{E}_{n+1} &= \big(1 - \tfrac{\sigma\tau}{4\varepsilon}\big)\mathbf{E}_{n+1/2} + \tfrac{\tau}{2\varepsilon} C_1 \mathbf{H}_{n+1/2} - \tfrac{\tau}{2\varepsilon} C_2 \mathbf{H}_{n+1} \\
&\quad - \big(1 - \tfrac{\sigma\tau}{4\varepsilon}\big)\tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1})), \\
\mathbf{H}_{n+1} &= \mathbf{H}_{n+1/2} + \tfrac{\tau}{2\mu} C_2 \mathbf{E}_{n+1/2} - \tfrac{\tau}{2\mu} C_1 \mathbf{E}_{n+1} \\
&\quad - \tfrac{\tau}{2\mu} C_2 \tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1})).
\end{aligned}$$

Plugging again the second equation into the first one gives in $L^2(Q)^3$ that

$$\begin{aligned}
\big((1 + \tfrac{\sigma\tau}{4\varepsilon})I - \tfrac{\tau^2}{4\varepsilon} D_\mu^{(2)}\big)\mathbf{E}_{n+1} &= \big(1 - \tfrac{\sigma\tau}{4\varepsilon}\big)\mathbf{E}_{n+1/2} + \tfrac{\tau}{2\varepsilon} \operatorname{curl} \mathbf{H}_{n+1/2} \\
&\quad - \tfrac{\tau^2}{4\varepsilon} C_2 \tfrac{1}{\mu} C_2 \mathbf{E}_{n+1/2} \\
&\quad - \big(1 - \tfrac{\sigma\tau}{4\varepsilon}\big)\tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1})) \\
&\quad + \tfrac{\tau^3}{8\varepsilon} C_2 \tfrac{1}{\mu} C_2 \tfrac{1}{\varepsilon}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1})), \\
\mathbf{H}_{n+1} &= \mathbf{H}_{n+1/2} + \tfrac{\tau}{2\mu} C_2 \mathbf{E}_{n+1/2} - \tfrac{\tau}{2\mu} C_1 \mathbf{E}_{n+1} \\
&\quad - \tfrac{\tau}{2\mu} C_2 \tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1})).
\end{aligned}$$

Also here we use the representation of $D_\mu^{(2)}$ as diagonal matrix in (8.11) to see that in this second half-step the implicit part of the first equation decouples into three independent equations, so that altogether the whole scheme can be computed efficiently.

9. Convergence of the ADI splitting scheme and preservation of the divergence conditions

In this chapter we investigate the convergence of the ADI scheme and the preservation of the divergence conditions of the numerical solutions. We treat both questions with respect to the L^2-norm and in a weaker sense (in Y^* and in $H^{-1}(Q)^6$, respectively).

In this chapter we assume without loss of generality that $\tau \leq 1$.

9.1. Convergence of the numerical scheme in L^2

The goal of this section is to prove the convergence of the ADI scheme (8.8) in $L^2(Q)$.

We integrate the convolution of the semigroup generated by the Maxwell operator with a polynomial and show afterwards some properties of the resulting operators. Recall Proposition 7.21 for the generation properties of $M_{\mathrm{div}}^{(0)}$ and $M_{\mathrm{div}}^{(2)}$.

Definition 9.1. *(a) We denote the C_0-semigroup generated by M by at time $\tau > 0$ by e^{tM} and define the operators $\Lambda_j(\tau)$ by*

$$\Lambda_j(\tau) := \frac{1}{\tau^j} \int_0^\tau e^{(\tau-s)M} \frac{s^{j-1}}{(j-1)!} \, \mathrm{d}s$$

for $j \geq 1$ and $\Lambda_0(\tau) := e^{\tau M}$.

(b) We denote the C_0-semigroup generated by $M_{\mathrm{div}}^{(0)}$ at time $\tau > 0$ by $e^{\tau M_{\mathrm{div}}^{(0)}}$ and define the operators $\Lambda_j^{(0)}(\tau)$ on $X_{\mathrm{div}}^{(0)}$ by

$$\Lambda_j^{(0)}(\tau) := \frac{1}{\tau^j} \int_0^\tau e^{(\tau-s)M_{\mathrm{div}}^{(0)}} \frac{s^{j-1}}{(j-1)!} \, \mathrm{d}s$$

for $j \geq 1$ and $\Lambda_0^{(0)}(\tau) := e^{\tau M_{\mathrm{div}}^{(0)}}$.

(c) If $\varepsilon, \mu, \sigma \in W^{2,3}(Q)$, then we denote the C_0-semigroup generated by $M_{\mathrm{div}}^{(2)}$ at time $\tau > 0$ by $e^{\tau M_{\mathrm{div}}^{(2)}}$ and define the operators $\Lambda_j^{(2)}(\tau)$ on $X_{\mathrm{div}}^{(2)}$ by

$$\Lambda_j^{(2)}(\tau) := \frac{1}{\tau^j} \int_0^\tau e^{(\tau-s)M_{\mathrm{div}}^{(2)}} \frac{s^{j-1}}{(j-1)!} \, \mathrm{d}s$$

for $j \geq 1$ and $\Lambda_0^{(2)}(\tau) := e^{\tau M_{\mathrm{div}}^{(2)}}$.

Lemma 9.2. *(a) For all $j \geq 0$ we have*

$$\Lambda_j(\tau) = \frac{1}{j!} I + \tau M \Lambda_{j+1}(\tau) \quad \text{on } X, \tag{9.1a}$$

$$\Lambda_j^{(0)}(\tau) = \frac{1}{j!} I + \tau M_{\mathrm{div}}^{(0)} \Lambda_{j+1}^{(0)}(\tau) \quad \text{on } X_{\mathrm{div}}^{(0)}, \tag{9.1b}$$

$$\Lambda_j^{(2)}(\tau) = \frac{1}{j!} I + \tau M_{\mathrm{div}}^{(2)} \Lambda_{j+1}^{(2)}(\tau) \quad \text{on } X_{\mathrm{div}}^{(2)}. \tag{9.1c}$$

(b) Under the assumption $\tau \leq 0$ we have

$$\|\Lambda_j(\tau)\|_X \leq \frac{C}{j!}, \quad \left\|\Lambda_j^{(0)}(\tau)\right\|_{X_{\mathrm{div}}^{(0)}} \leq \frac{C_0}{j!} \quad \text{and} \quad \left\|\Lambda_j^{(2)}(\tau)\right\|_{X_{\mathrm{div}}^{(2)}} \leq \frac{C_2}{j!},$$

with

$$C = \sup_{s\in[0,1]} \left\|e^{sM}\right\|_X, \quad C_0 = \sup_{s\in[0,1]} \left\|e^{sM_{\mathrm{div}}^{(0)}}\right\|_{X_{\mathrm{div}}^{(0)}} \quad \text{and} \quad C_2 = \sup_{s\in[0,1]} \left\|e^{sM_{\mathrm{div}}^{(2)}}\right\|_{X_{\mathrm{div}}^{(2)}}.$$

(c) For all $j \geq 0$ the operators $\Lambda_j(\tau)$, $\Lambda_j^{(0)}(\tau)$ and $\Lambda_j^{(2)}(\tau)$ leave $D(M)$, $D(M_{\mathrm{div}}^{(0)})$ and $D(M_{\mathrm{div}}^{(2)})$ invariant, respectively.

(d) For all $j \geq 1$ the operators $\Lambda_j(\tau)$, $\Lambda_j^{(0)}(\tau)$ and $\Lambda_j^{(2)}(\tau)$ map into $D(M)$, $D(M_{\mathrm{div}}^{(0)})$ and $D(M_{\mathrm{div}}^{(2)})$, respectively.

Proof:
(a) is seen with integration by parts. The rest of the statements follow easily from Definition 9.1, semigroup theory and Proposition 7.21. □

We are now in position to formulate and prove our first convergence theorem. Keep in mind that $D(M_{\mathrm{div}}^{(0)}) \hookrightarrow D(A) \cap D(B)$ by Proposition 7.15.

Theorem 9.3. *Let $T > 0$, $\varepsilon, \mu, \sigma \in W^{2,3}(Q)$, $(\boldsymbol{E}_0, \boldsymbol{H}_0) \in D(M_{\mathrm{div}}^{(2)})$ and*

$$\left(\tfrac{1}{\varepsilon}\boldsymbol{J}_0, 0\right) \in C^1\big([0,T], X_{\mathrm{div}}^{(2)}\big) \cap C^2\big([0,T], D(M_{\mathrm{div}}^{(0)})\big).$$

Then the numerical scheme (8.8) converges quadratically in $L^2(Q)^6$ to the solution of (7.1), i.e. for all $\tau > 0$ and $n \in \mathbb{N}$ with $n\tau \leq T$ we have

$$\begin{aligned}
&\left\|S_{\tau,n}^I \cdots S_{\tau,1}^I(\boldsymbol{E}_0, \boldsymbol{H}_0) - (\boldsymbol{E}(n\tau), \boldsymbol{H}(n\tau))\right\|_{L^2} \\
&\quad \leq C\tau^2 \left(T\Big(\|(\boldsymbol{E}_0, \boldsymbol{H}_0)\|_{D(M_{\mathrm{div}}^{(2)})} + \left\|\left(\tfrac{1}{\varepsilon}\boldsymbol{J}_0, 0\right)\right\|_{C^1\left([0,T], X_{\mathrm{div}}^{(2)}\right)} \Big) + \int_0^{n\tau} \|(\boldsymbol{J}_0''(s), 0)\|_{D(M_{\mathrm{div}}^{(0)})} \, \mathrm{d}s \right)
\end{aligned}$$

with a constant C only depending on $\|\varepsilon\|_{W^{1,\infty}\cap W^{2,3}}$, $\|\mu\|_{W^{1,\infty}\cap W^{2,3}}$, $\|\sigma\|_{W^{1,\infty}\cap W^{2,3}}$ and δ.

Remark 9.4. *It the solution is in $C\big([0,T], D(M^{(0)}_{\mathrm{div}})\big)$ with norm smaller than M, then it is sufficient to assume*

$$\big(\tfrac{1}{\varepsilon}\boldsymbol{J}_0, 0\big) \in C\big([0,T], X^{(2)}_{\mathrm{div}}\big) \cap C^2\big([0,T], D(M^{(0)}_{\mathrm{div}})\big).$$

In this case we have

$$\begin{aligned}
&\big\|S^I_{\tau,n}\cdots S^I_{\tau,1}(\boldsymbol{E}_0,\boldsymbol{H}_0) - (\boldsymbol{E}(n\tau),\boldsymbol{H}(n\tau))\big\|_{L^2}\\
&\quad\le c\tau^2\bigg(T\Big(M + \big\|\big(\tfrac{1}{\varepsilon}\boldsymbol{J}_0,0\big)\big\|_{C\left([0,T],X^{(2)}_{\mathrm{div}}\right)} + \|(\boldsymbol{J}_0',0)\|_{C^1\left([0,T],D(M^{(0)}_{\mathrm{div}})\right)}\Big)\\
&\qquad\qquad + \int_0^{n\tau} \|(\boldsymbol{J}_0''(s),0)\|_{D(M^{(0)}_{\mathrm{div}})}\,\mathrm{d}s\bigg).
\end{aligned}$$

Proof:
First observe that the embedding $X^{(2)}_{\mathrm{div}} \hookrightarrow D(A)\cap D(B)$ from Proposition 7.15 ensures that $S^I_{\tau,n}\cdots S^I_{\tau,1}(\mathbf{E}_0,\mathbf{H}_0)$ is well-defined for all $n\in\mathbb{N}$. Let $\tau>0$ and $n\in\mathbb{N}$ with $(n+1)\tau\le T$ be fixed. A Taylor expansion of $\mathbf{J}_0(n\tau+s)$ at $n\tau$ for $s\in(0,\tau]$ yields the identity

$$\big(\tfrac{1}{\varepsilon}\mathbf{J}_0(n\tau+s),0\big) = \bigg(\tfrac{1}{\varepsilon}\mathbf{J}_0(n\tau) + s\tfrac{1}{\varepsilon}\mathbf{J}_0'(n\tau) + \int_{n\tau}^{n\tau+s}(n\tau+s-r)\tfrac{1}{\varepsilon}\mathbf{J}_0''(r)\,\mathrm{d}r, 0\bigg) \tag{9.2}$$

in $X^{(2)}_{\mathrm{div}}$. By Theorem 1.9, the solution $w=(\mathbf{E},\mathbf{H})$ of (7.1) belongs to $C\big([0,T],D(M^{(2)}_{\mathrm{div}})\big)$ and can be written in $D(M^{(2)}_{\mathrm{div}})$, using (9.2) and Definition 9.1, as

$$\begin{aligned}
w((n+1)\tau) &= e^{\tau M}w(n\tau) + \int_0^\tau e^{(\tau-s)M}\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0(n\tau+s),0\big)\,\mathrm{d}s\\
&= e^{\tau M}w(n\tau) + \int_0^\tau e^{(\tau-s)M}\bigg(-\tfrac{1}{\varepsilon}\Big(\mathbf{J}_0(n\tau) + s\mathbf{J}_0'(n\tau)\\
&\quad + \int_{n\tau}^{n\tau+s}(n\tau+s-r)\mathbf{J}_0''(r)\,\mathrm{d}r\Big),0\bigg)\,\mathrm{d}s\\
&= \Lambda_0(\tau)w(n\tau) + \tau\Lambda_1(\tau)\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0(n\tau),0\big) + \tau^2\Lambda_2(\tau)\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0'(n\tau),0\big)\\
&\quad + R_n(\tau)
\end{aligned} \tag{9.3}$$

with

$$R_n(\tau) := \int_0^\tau e^{(\tau-s)M}\left(\int_{n\tau}^{n\tau+s}(n\tau+s-r)\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0''(r),0\big)\,\mathrm{d}r\right)\mathrm{d}s.$$

We have

$$\begin{aligned}
\big\|(I+\tfrac{\tau}{2}B)R_n(\tau)\big\|_X \le c\,\|R_n(\tau)\|_{D(M^{(0)}_{\mathrm{div}})} &\le c\tau^2\int_{n\tau}^{(n+1)\tau}\|(\mathbf{J}_0''(s),0)\|_{D(M^{(0)}_{\mathrm{div}})}\,\mathrm{d}s,\\
\|R_n(\tau)\|_X &\le c\tau^2\int_{n\tau}^{(n+1)\tau}\|(\mathbf{J}_0''(s),0)\|_{D(M^{(0)}_{\mathrm{div}})}\,\mathrm{d}s,
\end{aligned}$$

with the constants c only depending on $\|\varepsilon\|_{W^{1,\infty}}$, $\|\mu\|_{W^{1,\infty}}$, $\|\sigma\|_{W^{1,\infty}}$ and δ.

Plugging the Taylor expansion (9.2) with $s = \tau$ into the numerical scheme $S^I_{\tau,n+1}$ from (8.8) and applying it to $S^I_{\tau,n} \cdots S^I_{\tau,1} w(0)$ gives for all $n \in \mathbb{N}_0$ with $(n+1)\tau \le T$ in X that

$$\begin{aligned} S^I_{\tau,n+1} S^I_{\tau,n} \cdots S^I_{\tau,1} w(0) = (I - \tfrac{\tau}{2}B)^{-1}(I + \tfrac{\tau}{2}A)\Big[(I - \tfrac{\tau}{2}A)^{-1}(I + \tfrac{\tau}{2}B) S^I_{\tau,n} \cdots S^I_{\tau,1} w(0) \\ + \tau\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0(n\tau), 0\big) + \tfrac{1}{2}\tau^2\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0'(n\tau), 0\big)\Big] \\ + (I - \tfrac{\tau}{2}B)^{-1}(I + \tfrac{\tau}{2}A) r_n(\tau) \end{aligned} \tag{9.4}$$

with

$$r_n(\tau) := \frac{\tau}{2} \int_{n\tau}^{(n+1)\tau} ((n+1)\tau - r)\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0''(r), 0\big)\, \mathrm{d}r.$$

The assumption $\big(\mathbf{J}_0, 0\big) \in C^2\big([0,T], D(A)\big)$ implies $r_n \in C\big([0,T], D(A)\big)$ and hence, similar to above,

$$\big\|(I + \tfrac{\tau}{2}A) r_n(\tau)\big\|_X \le c\tau^2 \int_{n\tau}^{(n+1)\tau} \|(\mathbf{J}_0''(r), 0)\|_{D(A)}\, \mathrm{d}r$$

with c only depending on $\|\varepsilon\|_{W^{1,\infty}}$, $\|\mu\|_{W^{1,\infty}}$, $\|\sigma\|_{W^{1,\infty}}$ and δ.

We use the notation

$$\gamma_{\tau/2}(A) = (I + \tfrac{\tau}{2}A)(I - \tfrac{\tau}{2}A)^{-1}$$

and analogously for B instead of A. Taking the difference between (9.3) and (9.4), and using the embedding $X^{(2)}_{\mathrm{div}} \hookrightarrow D(AB) \cap D(A^2)$ from Proposition 7.15 and that $(I - \frac{\tau}{2}A)^{-1}$ and $I + \frac{\tau}{2}A$ commute on $D(A)$, we have for all $n \in \mathbb{N}_0$ with $(n+1)\tau \le T$ that

$$\begin{aligned} &S^I_{\tau,n+1} S^I_{\tau,n} \cdots S^I_{\tau,1} w(0) - w((n+1)\tau) \\ &\quad = (I - \tfrac{\tau}{2}B)^{-1}\gamma_{\tau/2}(A)(I + \tfrac{\tau}{2}B)\big(S^I_{\tau,n} \cdots S^I_{\tau,1} w(0) - w(n\tau)\big) \\ &\qquad + (I - \tfrac{\tau}{2}B)^{-1}(I - \tfrac{\tau}{2}A)^{-1} \cdot \\ &\qquad\quad \cdot \big((I + \tfrac{\tau}{2}A)(I + \tfrac{\tau}{2}B) - (I - \tfrac{\tau}{2}A)(I - \tfrac{\tau}{2}B)\Lambda_0^{(2)}(\tau)\big) w(n\tau) \\ &\qquad + \tau(I - \tfrac{\tau}{2}B)^{-1}(I - \tfrac{\tau}{2}A)^{-1} \cdot \\ &\qquad\quad \cdot \big((I - \tfrac{\tau}{2}A)(I + \tfrac{\tau}{2}A) - (I - \tfrac{\tau}{2}A)(I - \tfrac{\tau}{2}B)\Lambda_1^{(2)}(\tau)\big)\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0(n\tau), 0\big) \\ &\qquad + \tau^2(I - \tfrac{\tau}{2}B)^{-1}\big(\tfrac{1}{2}(I + \tfrac{\tau}{2}A) - (I - \tfrac{\tau}{2}B)\Lambda_2^{(0)}(\tau)\big)\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0'(n\tau), 0\big) \\ &\qquad + (I - \tfrac{\tau}{2}B)^{-1}(I + \tfrac{\tau}{2}A) r_n(\tau) - R_n(\tau) \\ &\quad =: (I - \tfrac{\tau}{2}B)^{-1}\gamma_{\tau/2}(A)(I + \tfrac{\tau}{2}B)\big(S^I_{\tau,n} \cdots S^I_{\tau,1} w(0) - w(n\tau)\big) \\ &\qquad + \Sigma_1(\tau) + \Sigma_2(\tau) + \Sigma_3(\tau) + (I - \tfrac{\tau}{2}B)^{-1}(I + \tfrac{\tau}{2}A) r_n(\tau) - R_n(\tau). \end{aligned}$$

Using (9.1), we see as in Section 4.1 of [37] that

$$\begin{aligned} \Sigma_1(\tau) = \tau^3 (I - \tfrac{\tau}{2}B)^{-1}(I - \tfrac{\tau}{2}A)^{-1}\Big(\big(M^{(0)}_{\mathrm{div}}\big)^2\big(\tfrac{1}{2}\Lambda_2^{(2)}(\tau) - \Lambda_3^{(2)}(\tau)\big) M^{(2)}_{\mathrm{div}} \\ - \tfrac{1}{4} AB \Lambda_1^{(2)}(\tau) M^{(2)}_{\mathrm{div}}\Big) w(n\tau). \end{aligned}$$

We recall that $M = A + B$ on $X^{(2)}_{\mathrm{div}}$ and the embedding $X^{(2)}_{\mathrm{div}} \hookrightarrow D\big((M^{(0)}_{\mathrm{div}})^2\big)$ from Coroallary 7.13. Taking this and (9.1) into account, we have in X^A_{-1} the identity

$$\begin{aligned}
&(I - \tfrac{\tau}{2}A)(I + \tfrac{\tau}{2}A) - (I - \tfrac{\tau}{2}A)(I - \tfrac{\tau}{2}B)\Lambda^{(2)}_1(\tau) \\
&\quad = I - \tfrac{\tau^2}{4}A^2 - (I - \tfrac{\tau}{2}(A+B) + \tfrac{\tau^2}{4}A_{-1}B)\big(I + \tau\Lambda^{(0)}_2(\tau)M^{(0)}_{\mathrm{div}}\big) \\
&\quad = -\tfrac{\tau^2}{4}A^2 - \tau M^{(0)}_{\mathrm{div}}\big(\tfrac{1}{2}I + \tau\Lambda^{(0)}_3(\tau)M^{(0)}_{\mathrm{div}}\big) + \tfrac{\tau}{2}M^{(0)}_{\mathrm{div}} + \tfrac{\tau^2}{2}\Lambda^{(0)}_2(\tau)\big(M^{(0)}_{\mathrm{div}}\big)^2 \\
&\qquad - \tfrac{\tau^2}{4}AB - \tfrac{\tau^3}{4}A_{-1}B\Lambda^{(0)}_2(\tau)M^{(0)}_{\mathrm{div}} \\
&\quad = -\tfrac{\tau^2}{4}A^2 - \tau^2\Lambda^{(0)}_3(\tau)\big(M^{(0)}_{\mathrm{div}}\big)^2 + \tfrac{\tau^2}{2}\Lambda^{(0)}_2(\tau)\big(M^{(0)}_{\mathrm{div}}\big)^2 - \tfrac{\tau^2}{4}AB - \tfrac{\tau^2}{4}AB(\Lambda^{(2)}_1(\tau) - I)
\end{aligned}$$

of operators acting on $X^{(2)}_{\mathrm{div}}$. Thus,

$$\begin{aligned}
\Sigma_2(\tau) = \tau^3(I - \tfrac{\tau}{2}B)^{-1}(I - \tfrac{\tau}{2}A)^{-1}\Big(&-\tfrac{1}{4}A^2 - \Lambda^{(0)}_3(\tau)\big(M^{(0)}_{\mathrm{div}}\big)^2 \\
&+ \tfrac{1}{2}\Lambda^{(0)}_2(\tau)\big(M^{(0)}_{\mathrm{div}}\big)^2 - \tfrac{1}{4}AB - \tfrac{1}{4}AB(\Lambda^{(2)}_2(\tau) - I)\Big)\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0(n\tau), 0\big).
\end{aligned}$$

Next, we conclude by (9.1) and $D(M^{(0)}_{\mathrm{div}}) \hookrightarrow D(A) \cap D(B)$ from Proposition 7.15 the identity

$$\begin{aligned}
\tfrac{1}{2}(I + \tfrac{\tau}{2}A) - (I - \tfrac{\tau}{2}B)\Lambda^{(0)}_2(\tau) &= \tfrac{1}{2}I + \tfrac{\tau}{4}A - (I - \tfrac{\tau}{2}B)\big(\tfrac{1}{2}I + \tau\Lambda^{(0)}_3(\tau)M^{(0)}_{\mathrm{div}}\big) \\
&= \tfrac{\tau}{4}A + \tfrac{\tau}{4}B - \tau\Lambda^{(0)}_3(\tau)M^{(0)}_{\mathrm{div}} + \tfrac{\tau}{2}B(\Lambda^{(0)}_2(\tau) - I)
\end{aligned}$$

on $D(M^{(0)}_{\mathrm{div}})$. This implies

$$\Sigma_3(\tau) = \tau^3(I - \tfrac{\tau}{2}B)^{-1}\big(\tfrac{1}{4}A - \tfrac{1}{4}B - \Lambda^{(0)}_3(\tau)M^{(0)}_{\mathrm{div}} + \tfrac{\tau}{2}\Lambda^{(0)}_3M^{(0)}_{\mathrm{div}}\big)\big(-\tfrac{1}{\varepsilon}\mathbf{J}'_0(n\tau), 0\big).$$

We abbreviate

$$\begin{aligned}
J_k(\tau) :&= (I - \tfrac{\tau}{2}A)^{-1}\Big(\big(M^{(0)}_{\mathrm{div}}\big)^2\big(\tfrac{1}{2}\Lambda^{(2)}_2(\tau) - \Lambda^{(2)}_3(\tau)\big)M^{(2)}_{\mathrm{div}} - \tfrac{1}{4}AB\Lambda^{(2)}_1(\tau)M^{(2)}_{\mathrm{div}}\Big)w(k\tau) \\
&\quad + (I - \tfrac{\tau}{2}A)^{-1}\Big(-\tfrac{1}{4}A^2 - \Lambda^{(0)}_3(\tau)\big(M^{(0)}_{\mathrm{div}}\big)^2 \\
&\qquad + \tfrac{1}{2}\Lambda^{(0)}_2(\tau)\big(M^{(0)}_{\mathrm{div}}\big)^2 - \tfrac{1}{4}AB - \tfrac{1}{4}AB(\Lambda^{(2)}_2(\tau) - I)\Big)\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0(k\tau), 0\big) \\
&\quad + \Big(\tfrac{1}{4}A - \tfrac{1}{4}B - \Lambda^{(0)}_3(\tau)M^{(0)}_{\mathrm{div}} + \tfrac{\tau}{2}\Lambda^{(0)}_3(\tau)M^{(0)}_{\mathrm{div}}\Big)\big(-\tfrac{1}{\varepsilon}\mathbf{J}'_0(k\tau), 0\big)
\end{aligned}$$

for all $k \geq 0$ with $k\tau \leq T$. We can estimate this expression by

$$\|J_k(\tau)\|_{L^2} \leq c\Big(\|w(k\tau)\|_{D(M^{(2)}_{\mathrm{div}})} + \big\|\big(\tfrac{1}{\varepsilon}\mathbf{J}_0(k\tau), 0\big)\big\|_{X^{(2)}_{\mathrm{div}}} + \|(\mathbf{J}'_0(k\tau), 0)\|_{D(M^{(0)}_{\mathrm{div}})}\Big)$$

with c only depending on $\|\varepsilon\|_{W^{1,\infty}\cap W^{2,3}}$, $\|\mu\|_{W^{1,\infty}\cap W^{2,3}}$, $\|\sigma\|_{W^{1,\infty}\cap W^{2,3}}$, δ and T, see Proposition 8.1, 7.11 and 7.15, and Lemma 9.2. The above calculations yield

$$S^I_{\tau,n}\cdots S^I_{\tau,1}w(0) - w(n\tau)$$

$$
\begin{aligned}
&= (I - \tfrac{\tau}{2}B)^{-1}\gamma_{\tau/2}(A)(I + \tfrac{\tau}{2}B)\big(S^I_{\tau,n-1}\cdots S^I_{\tau,1}w(0) - w((n-1)\tau)\big)\\
&\quad + \tau^3 J_{n-1}(\tau) + (I - \tfrac{\tau}{2}B)^{-1}(I + \tfrac{\tau}{2}A)r_{n-1}(\tau) - R_{n-1}(\tau).
\end{aligned}
$$

We solve this error recursion and get

$$
\begin{aligned}
&S^I_{\tau,n}\cdots S^I_{\tau,1}w(0) - w(n\tau)\\
&\quad = \tau^3\sum_{k=0}^{n-1}(I - \tfrac{\tau}{2}B)^{-1}\big(\gamma_{\tau/2}(A)\gamma_{\tau/2}(B)\big)^{n-1-k}J_k(\tau)\\
&\qquad + \sum_{k=0}^{n-1}(I - \tfrac{\tau}{2}B)^{-1}\big(\gamma_{\tau/2}(A)\gamma_{\tau/2}(B)\big)^{n-1-k}(I + \tfrac{\tau}{2}A)r_k(\tau)\\
&\qquad - \sum_{k=0}^{n-2}(I - \tfrac{\tau}{2}B)^{-1}\big(\gamma_{\tau/2}(A)\gamma_{\tau/2}(B)\big)^{n-2-k}\gamma_{\tau/2}(A)(I + \tfrac{\tau}{2}B)R_k(\tau)\\
&\qquad - R_{n-1}(\tau).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&\big\|S^I_{\tau,n}\cdots S^I_{\tau,1}w_0 - w(n\tau)\big\|_{L^2}\\
&\le c\tau^3\sum_{k=0}^{n-1}\Bigg(\|w(k\tau)\|_{D(M^{(2)}_{\mathrm{div}})} + \big\|\big(\tfrac{1}{\varepsilon}\mathbf{J}_0(k\tau),0\big)\big\|_{X^{(2)}_{\mathrm{div}}} + \|(\mathbf{J}_0'(k\tau),0)\|_{D(M^{(0)}_{\mathrm{div}})}\\
&\qquad\qquad + c\tau^2\int_{k\tau}^{(k+1)\tau}\|(\mathbf{J}_0''(r),0)\|_{D(M^{(0)}_{\mathrm{div}})}\,\mathrm{d}r\Bigg)\\
&\le C\tau^2\Bigg(T\Big(\|w_0\|_{D(M^{(2)}_{\mathrm{div}})} + \big\|\big(\tfrac{1}{\varepsilon}\mathbf{J}_0,0\big)\big\|_{C^1([0,T],X^{(2)}_{\mathrm{div}})}\Big) + \int_0^{n\tau}\|(\mathbf{J}_0''(r),0)\|_{D(M^{(0)}_{\mathrm{div}})}\,\mathrm{d}r\Bigg),
\end{aligned}
$$

see Proposition 8.1 and Theorem 1.9. Thereby, C only depends on $\|\varepsilon\|_{W^{1,\infty}\cap W^{2,3}}$, $\|\mu\|_{W^{1,\infty}\cap W^{2,3}}$, $\|\sigma\|_{W^{1,\infty}\cap W^{2,3}}$ and δ. □

9.2. Convergence of the numerical scheme in a weak sense

We first remind the reader that for $w_0 = (\mathbf{E}_0, \mathbf{H}_0) \in D(M^{(0)}_{\mathrm{div}})$ and

$$
\big(\mathbf{J}_0, 0\big) \in C\big([0,T], D(M^{(0)}_{\mathrm{div}})\big) + C^1\big([0,T], X^{(0)}_{\mathrm{div}}\big),
$$

Proposition 7.21 gives a unique solution w of (7.1) with

$$
w = (\mathbf{E}, \mathbf{H}) \cap C^1\big([0,T], X^{(0)}_{\mathrm{div}}\big) \cap C\big([0,T], D(M^{(0)}_{\mathrm{div}})\big).
$$

Moreover, we recall the definition of Sobolev spaces of negative orders associated to semigroups from Proposition 1.10.

With reduced regularity assumptions on the initial function and on the inhomogeneity our numerical scheme is still convergent of order two in time in a weak sense.

Theorem 9.5. *Let $T > 0$, $(\boldsymbol{E}_0, \boldsymbol{H}_0) \in D\big((M^{(0)}_{\mathrm{div}})^2\big)$ and*

$$\big(\boldsymbol{J}_0, 0\big) \in C\big([0,T], D\big(M^{(0)}_{\mathrm{div}}\big)\big) \cap C^2\big([0,T], X^{(0)}_{\mathrm{div}}\big).$$

Then the numerical scheme (8.8) *converges for small time step sizes quadratically in Y^* to the solution of* (7.1)*, i.e. there is a bound $\tau_0 \in [0,T)$ on the time step size such that for all $\tau \in (0,\tau_0]$ and $n \in \mathbb{N}$ with $n\tau \le T$ we have*

$$\begin{aligned}
&\big|\big(S^I_{\tau,n} \cdots S^I_{\tau,1}(\boldsymbol{E}_0, \boldsymbol{H}_0) - (\boldsymbol{E}(n\tau), \boldsymbol{H}(n\tau)) \mid (\varphi,\psi)\big)_X\big| \\
&\quad \le C\tau^2 e^{6\kappa_Y T} T\Big(\|(\boldsymbol{E}_0, \boldsymbol{H}_0)\|_{D\left((M^{(0)}_{\mathrm{div}})^2\right)} + \|(\boldsymbol{J}_0, 0)\|_{C\left([0,T], D(M^{(0)}_{\mathrm{div}})\right)} \\
&\qquad + \|(\boldsymbol{J}_0, 0)\|_{C^2\left([0,T], X^{(0)}_{\mathrm{div}}\right)}\Big) \|(\varphi,\psi)\|_Y
\end{aligned}$$

for all $(\varphi,\psi) \in Y$ with a constant C only depending on $\|\varepsilon\|_{W^{1,\infty}}$, $\|\mu\|_{W^{1,\infty}}$, $\|\sigma\|_{W^{1,\infty}}$ and δ.

Proof:
First observe that the embedding $D(M^{(0)}_{\mathrm{div}}) \hookrightarrow D(A) \cap D(B)$ from Proposition 7.15 ensures that $S^I_{\tau,n} \cdots S^I_{\tau,1}(\mathbf{E}_0, \mathbf{H}_0)$ is well-defined for all $n \in \mathbb{N}$.

Let $\tau_0 := \min\{\frac{1}{2\kappa_Y}, \widetilde{\tau}\}$ with the $\widetilde{\tau}$ from Proposition 8.7 and let $\tau \in (0,\tau_0]$. Let $n \in \mathbb{N}$ with $n\tau \le T$. Let $(\varphi,\psi) = (I + \frac{\tau}{2}B^*)(\varphi_0,\psi_0)$ for some $(\varphi_0,\psi_0) \in Y$. Under the regularity assumptions of this theorem the Taylor expansion (9.2) is only valid in X and (9.3) is valid in X with

$$R_n(\tau) := \int_0^\tau e^{(\tau-s)M^{(0)}_{\mathrm{div}}} \left(\int_{n\tau}^{n\tau+s} (n\tau + s - r)\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0''(r), 0\big)\,\mathrm{d}r \right) \mathrm{d}s.$$

We get, due to $Y \hookrightarrow D(B^*)$ by Remark 7.16, that

$$\begin{aligned}
\big|\big(R_n(\tau) \mid (\varphi,\psi)\big)_X\big| &\le c \int_0^\tau \big\|e^{(\tau-s)M}\big\|_{\mathcal{B}(X)} \cdot \\
&\quad \cdot \int_{n\tau}^{n\tau+s} (n\tau + s - r)\,\big\|\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0''(r), 0\big)\big\|_X \,\mathrm{d}r\,\mathrm{d}s\,\|(\varphi,\psi)\|_X \\
&\le c\tau^2 \int_{n\tau}^{(n+1)\tau} \|(\mathbf{J}_0''(r), 0)\|_X \,\mathrm{d}r\,\|(\varphi,\psi)\|_X
\end{aligned}$$

and

$$\big|\big(R_n(\tau) \mid (\varphi,\psi)\big)_X\big| \le c\tau^2 \int_{n\tau}^{(n+1)\tau} \|(\mathbf{J}_0''(r), 0)\|_X \,\mathrm{d}r\,\|(\varphi,\psi)\|_X$$

with the constants c only depending on $\|\varepsilon\|_{L^\infty}$, $\|\mu\|_{L^\infty}$, $\|\sigma\|_{L^\infty}$ and δ. In the same way we get for

$$r_n(\tau) := \frac{\tau}{2} \int_{n\tau}^{(n+1)\tau} ((n+1)\tau - r)\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0''(r), 0\big)\,\mathrm{d}r$$

with $Y \hookrightarrow D(A^*)$ that

$$\big|\big(r_n(\tau) \mid (\varphi,\psi)\big)_X\big| \le c\tau^2 \int_{n\tau}^{(n+1)\tau} \|(\mathbf{J}_0''(r),0)\|_X \,\mathrm{d}r\, \|(\varphi,\psi)\|_X\,.$$

As in Section 9.1, we expand the inhomogeneity in both summands of

$$S^I_{\tau,n+1}\cdots S^I_{\tau,1}w(0) - w((n+1)\tau)$$

into a Taylor series, test the difference with (φ,ψ), bring the operators by taking the adjoints to the right-hand side and do the same algebraic reformulations as before. This gives

$$\begin{aligned}
&\big(S^I_{\tau,n+1}\cdots S^I_{\tau,1}w(0) - w((n+1)\tau) \mid (\varphi,\psi)\big)_X\\
&\quad= \big(S^I_{\tau,n}\cdots S^I_{\tau,1}w(0) - w(n\tau) \mid (I+\tfrac{\tau}{2}B^*)(I-\tfrac{\tau}{2}A^*)^{-1}(I+\tfrac{\tau}{2}A^*)(I-\tfrac{\tau}{2}B^*)^{-1}(\varphi,\psi)\big)_X\\
&\qquad+\Big(w(n\tau) \mid \big((I+\tfrac{\tau}{2}B^*)(I+\tfrac{\tau}{2}A^*)\\
&\qquad\qquad -\Lambda_0(\tau)^*(I-\tfrac{\tau}{2}B^*)(I-\tfrac{\tau}{2}A^*)\big)(I-\tfrac{\tau}{2}A^*)^{-1}(I-\tfrac{\tau}{2}B^*)^{-1}(\varphi,\psi)\Big)_X\\
&\qquad+\tau\Big(\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0(n\tau),0\big) \mid \big((I+\tfrac{\tau}{2}A^*)(I-\tfrac{\tau}{2}B^*)^{-1}-\Lambda_1(\tau)^*\big)(\varphi,\psi)\Big)_X\\
&\qquad+\tau^2\Big(\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0'(n\tau),0\big) \mid \big(\tfrac{1}{2}(I+\tfrac{\tau}{2}A^*)(I-\tfrac{\tau}{2}B^*)^{-1}-\Lambda_2(\tau)^*\big)(\varphi,\psi)\Big)_X\\
&\qquad+\big(r_n(\tau) \mid (I+\tfrac{\tau}{2}A^*)(I+\tfrac{\tau}{2}B^*)^{-1}(\varphi,\psi)\big)_{L^2} - (R_n(\tau)\mid(\varphi,\psi))_X\\
&\quad=:\big((I-\tfrac{\tau}{2}B)^{-1}(I+\tfrac{\tau}{2}A)(I-\tfrac{\tau}{2}A)^{-1}(I+\tfrac{\tau}{2}B)(S^I_{\tau,n}\cdots S^I_{\tau,1}w(0)-w(n\tau)) \mid (\varphi,\psi)\big)_X\\
&\qquad+\Sigma_1(\tau)+\Sigma_2(\tau)+\Sigma_3(\tau)+\big(r_n(\tau)\mid(I+\tfrac{\tau}{2}A^*)(I+\tfrac{\tau}{2}B^*)^{-1}(\varphi,\psi)\big)_X\\
&\qquad-(R_n(\tau)\mid(\varphi,\psi))_X,
\end{aligned}$$

where we used that $I+\frac{\tau}{2}A^*$ and $(I-\frac{\tau}{2}A^*)^{-1}$ commute on $Y \hookrightarrow D(A^*)\cap D(B^*)$. We set

$$\chi(\tau) := (I-\tfrac{\tau}{2}A_Y^*)^{-1}(I-\tfrac{\tau}{2}B^*)^{-1}(\varphi,\psi) \in D(A_Y^*)$$

and have due to $M^* = A^* + B^*$ on Y that

$$\begin{aligned}
\Sigma_1(\tau) &= \big(w(n\tau) \mid \big((I+\tfrac{\tau}{2}B^*)(I+\tfrac{\tau}{2}A^*) - \Lambda_0(\tau)^*(I-\tfrac{\tau}{2}B^*)(I-\tfrac{\tau}{2}A^*)\big)\chi(\tau)\big)_X\\
&= \Big(w(n\tau)\big|\big((I-\Lambda_0(\tau)^*) + \tfrac{\tau}{2}(I+\Lambda_0(\tau)^*)M^* + \tfrac{\tau^2}{4}(I-\Lambda_0(\tau)^*)B^*A^*\big)\chi(\tau)\Big)_X.
\end{aligned}$$

Due to (9.1) we have

$$\begin{aligned}
&I-\Lambda_0(\tau)^* = -\tau M^* - \tfrac{1}{2}\tau^2(M^*)^2 - \tau^3\Lambda_0(\tau)^*(M^*)^3 \quad \text{on } D\big((M^*)^3\big),\\
&I+\Lambda_0(\tau)^* = 2I + \tau M^* + \tau^2\Lambda_2(\tau)^*(M^*)^2 \quad \text{on } D\big((M^*)^2\big) \qquad \text{and}\\
&I-\Lambda_0(\tau)^* = -\tau\Lambda_1(\tau)^*M^* \quad \text{on } D(M^*).
\end{aligned}$$

Thus, using $Y \hookrightarrow D(B^*)$, $w(n\tau) \in D\big((M_{\mathrm{div}}^{(0)})^2\big) \hookrightarrow D(M^2)$ and $D(M_{\mathrm{div}}^{(0)}) \hookrightarrow D(A) \cap D(B)$ from Proposition 7.15, we get

$$\begin{aligned}
\Sigma_1(\tau) &= \Big\langle w(n\tau), \Big(-\tau^3 \Lambda_3(\tau)^*_{-2} M^*_{-2} M^*_{-1} M^* \\
&\qquad + \tfrac{\tau^3}{2} \Lambda_2(\tau)^*_{-2} M^*_{-2} M^*_{-1} M^* \\
&\qquad - \tfrac{\tau^3}{4} \Lambda_1(\tau)^*_{-1} M^*_{-1} B^* A^* \Big) \chi(\tau) \Big\rangle_{D(M^2) \times X^{M^*}_{-2}} \\
&= \tau^3 \Big(\big(-\big(M_{\mathrm{div}}^{(0)}\big)^2 \Lambda_3^{(0)}(\tau) + \tfrac{1}{2} (M_{\mathrm{div}}^{(0)})^2 \Lambda_2^{(0)}(\tau) \big) w(n\tau) \mid M_{\mathrm{div}}^{(0)} \chi(\tau) \Big)_X \\
&\quad - \tau^3 \big(\tfrac{1}{4} B M_{\mathrm{div}}^{(0)} \Lambda_1^{(0)}(\tau) w(n\tau) \mid A^* \chi(\tau) \big)_X .
\end{aligned}$$

Moreover,

$$\begin{aligned}
\Sigma_2(\tau) &= \tau \Big(\big(-\tfrac{1}{\varepsilon} \mathbf{J}_0(n\tau), 0 \big) \mid \big((I + \tfrac{\tau}{2} A^*)(I - \tfrac{\tau}{2} A^*) - \Lambda_1(\tau)^* (I - \tfrac{\tau}{2} B^*)(I - \tfrac{\tau}{2} A^*) \big) \chi(\tau) \Big)_X \\
&= \tau \Big(\big(-\tfrac{1}{\varepsilon} \mathbf{J}_0(n\tau), 0 \big) \mid \big(I - \tfrac{\tau^2}{4} (A^*)^2 - \Lambda_1(\tau)^* \big(I - \tfrac{\tau}{2} (A^* + B^*) \big) \\
&\qquad - \tfrac{\tau^2}{4} \Lambda_1(\tau)^* B^* A^* \big) \chi(\tau) \Big)_X .
\end{aligned}$$

Using first $\Lambda_1(\tau)^* = I + \tau \Lambda_2(\tau)^* M^*$ and then $\Lambda_2(\tau)^* = \frac{1}{2} I + \tau \Lambda_3(\tau)^* M^*$ by (9.1), we thus have

$$\begin{aligned}
\Sigma_2(\tau) &= \tau \Big\langle \big(-\tfrac{1}{\varepsilon} \mathbf{J}_0(n\tau), 0 \big), \big(-\tfrac{\tau^2}{4} (A^*)^2 - \tau \Lambda_2^{(0)}(\tau)^* M^*_{-1} + \tfrac{\tau}{2} M^* \\
&\qquad + \tfrac{\tau^2}{2} \Lambda_2(\tau)^*_{-1} M^*_{-1} M^* - \tau^2 \Lambda_1(\tau)^* B^* A^* \big) \chi(\tau) \Big\rangle_{D(M) \times X^{M^*}_{-1}} \\
&= \tau \Big\langle \big(-\tfrac{1}{\varepsilon} \mathbf{J}_0(n\tau), 0 \big), \big(-\tfrac{\tau^2}{4} (A^*)^2 - \tau^2 \Lambda_3(\tau)^*_{-1} M^*_{-1} M^* \\
&\qquad + \tfrac{\tau^2}{2} \Lambda_2(\tau)^*_{-1} M^*_{-1} M^* \\
&\qquad - \tau^2 \Lambda_1(\tau)^* B^* A^* \big) \chi(\tau) \Big\rangle_{D(M) \times X^{M^*}_{-1}} \\
&= \tau^3 \Big(\Big(-\tfrac{1}{4} A \big(-\tfrac{1}{\varepsilon} \mathbf{J}_0(n\tau), 0 \big) \mid A^* \chi(\tau) \Big)_X \\
&\qquad + \Big(-(M_{\mathrm{div}}^{(0)}) \Lambda_3^{(0)}(\tau) \big(-\tfrac{1}{\varepsilon} \mathbf{J}_0(n\tau), 0 \big) \mid M^* \chi(\tau) \Big)_X \\
&\qquad + \Big(-\tfrac{1}{2} M_{\mathrm{div}}^{(0)} \Lambda_2^{(0)}(\tau) \big(-\tfrac{1}{\varepsilon} \mathbf{J}_0(n\tau), 0 \big) \mid M^* \chi(\tau) \Big)_X \\
&\qquad + \Big(-\tfrac{1}{4} B \Lambda_1^{(0)}(\tau) \big(-\tfrac{1}{\varepsilon} \mathbf{J}_0(n\tau), 0 \big) \mid A^* \chi(\tau) \Big)_X \Big),
\end{aligned}$$

where we have taken $M_{\mathrm{div}}^{(0)} = A + B$ on $D(M_{\mathrm{div}}^{(0)})$ into account in the last equality. Furthermore, we have again with (9.1) that

$$\Sigma_3(\tau) = \tau^2 \Big(\big(-\tfrac{1}{\varepsilon} \mathbf{J}_0'(n\tau), 0 \big), \big(\tfrac{1}{2} (I + \tfrac{\tau}{2} A^*) - \Lambda_2(\tau)^* (I - \tfrac{\tau}{2} B^*) \big) (I - \tfrac{\tau}{2} B_Y^*)^{-1} (\varphi, \psi) \Big)_X$$

$$
\begin{aligned}
&= \tau^2\Big(\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0'(n\tau),0\big),\big(\tfrac{1}{2}I+\tfrac{\tau}{4}A^*-\big(\tfrac{1}{2}I+\tau\Lambda_3(\tau)^*M^*\big)\\
&\qquad+\tfrac{1}{2}\Lambda_2^{(0)}(\tau)^*B^*\big)(I-\tfrac{\tau}{2}B_Y^*)^{-1}(\varphi,\psi)\Big)_X\\
&= \tau^3\Big(\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0'(n\tau),0\big),\big(\tfrac{1}{4}A^*-\Lambda_3(\tau)^*M^*\\
&\qquad+\tfrac{1}{2}\Lambda_2(\tau)^*B^*\big)(I-\tfrac{\tau}{2}B_Y^*)^{-1}(\varphi,\psi)\Big)_X.
\end{aligned}
$$

We altogether get

$$
\begin{aligned}
&\big(S_{\tau,n}^I\cdots S_{\tau,1}^I w(0)-w(n\tau)\mid(\varphi,\psi)\big)_X\\
&\quad=\tau^3\sum_{k=0}^{n-1}\Big(\big(-(M_{\mathrm{div}}^{(0)})^2\Lambda_3^{(0)}(\tau)+\tfrac{1}{2}(M_{\mathrm{div}}^{(0)})^2\Lambda_2^{(0)}(\tau)\big)w(k\tau)\mid\\
&\qquad M^*(I-\tfrac{\tau}{2}A^*)^{-1}\big(\gamma_{\tau/2}(B)^*\gamma_{\tau/2}(A)^*\big)^{n-1-k}(I-\tfrac{\tau}{2}B^*)^{-1}(\varphi,\psi)\Big)_X\\
&\quad-\tau^3\sum_{k=0}^{n-1}\Big(\tfrac{1}{4}BM_{\mathrm{div}}^{(0)}\Lambda_1^{(0)}(\tau)w(k\tau)\mid A^*(I-\tfrac{\tau}{2}A^*)^{-1}\big(\gamma_{\tau/2}(B)^*\gamma_{\tau/2}(A)^*\big)^{n-1-k}\cdot\\
&\qquad(I+\tfrac{\tau}{2}B^*)^{-1}(\varphi,\psi)\Big)_X\\
&\quad-\tau^3\sum_{k=0}^{n-1}\Big(\big(\tfrac{1}{4}A+\tfrac{1}{4}B\Lambda_1^{(0)}(\tau)\big)\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0(k\tau),0\big)\mid\\
&\qquad A^*(I-\tfrac{\tau}{2}A^*)^{-1}\big(\gamma_{\tau/2}(B)^*\gamma_{\tau/2}(A)^*\big)^{n-1-k}(I+\tfrac{\tau}{2}B^*)^{-1}(\varphi,\psi)\Big)_X\\
&\quad-\tau^3\sum_{k=0}^{n-1}\Big(\tfrac{1}{2}M_{\mathrm{div}}^{(0)}\Lambda_2^{(0)}(\tau)\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0(k\tau),0\big)\mid M^*(I-\tfrac{\tau}{2}A^*)^{-1}\cdot\\
&\qquad\big(\gamma_{\tau/2}(B)^*\gamma_{\tau/2}(A)^*\big)^{n-1-k}(I-\tfrac{\tau}{2}B^*)^{-1}(\varphi,\psi)\Big)_X\\
&\quad-\tau^3\sum_{k=0}^{n-1}\Big(M_{\mathrm{div}}^{(0)}\Lambda_3^{(0)}(\tau)\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0(k\tau),0\big)\mid M^*(I-\tfrac{\tau}{2}A^*)^{-1}\cdot\\
&\qquad\big(\gamma_{\tau/2}(B)^*\gamma_{\tau/2}(A)^*\big)^{n-1-k}(I-\tfrac{\tau}{2}B^*)^{-1}(\varphi,\psi)\Big)_X\\
&\quad+\tau^3\sum_{k=0}^{n-1}\Big(\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0'(k\tau),0\big)\mid-\Lambda_3(\tau)^*M^*\big(\gamma_{\tau/2}(B)^*\gamma_{\tau/2}(A)^*\big)^{n-1-k}(I-\tfrac{\tau}{2}B^*)^{-1}(\varphi,\psi)\Big)_X\\
&\quad+\tau^3\sum_{k=0}^{n-1}\Big(\big(-\tfrac{1}{\varepsilon}\mathbf{J}_0'(k\tau),0\big)\mid\big(\tfrac{1}{4}A^*+\tfrac{1}{2}\Lambda_2^{(0)}(\tau)^*B^*\big)\big(\gamma_{\tau/2}(B)^*\gamma_{\tau/2}(A)^*\big)^{n-1-k}\cdot\\
&\qquad(I-\tfrac{\tau}{2}B^*)^{-1}(\varphi,\psi)\Big)_X\\
&\quad+\sum_{k=0}^{n-1}\Big(r_k(\tau)\mid(I+\tfrac{\tau}{2}A^*)\big(\gamma_{\tau/2}(B)^*\gamma_{\tau/2}(A)^*\big)^{n-1-k}(I-\tfrac{\tau}{2}B^*)^{-1}(\varphi,\psi)\Big)_X
\end{aligned}
$$

$$+\sum_{k=0}^{n-2}\Big(R_k(\tau)\mid (I+\tfrac{\tau}{2}B^*)\gamma_{\tau/2}(A)^*\big(\gamma_{\tau/2}(B)^*\gamma_{\tau/2}(A)^*\big)^{n-2-k}(I-\tfrac{\tau}{2}B^*)^{-1}(\varphi,\psi)\Big)_X$$
$$+\big(R_{n-1}(\tau)\mid(\varphi,\psi)\big)_X.$$

Analogously as in Section 9.1 we use the norm estimates from Proposition 8.1 and Theorem 1.9 to infer

$$\begin{aligned}
&\big|\big(S^I_{\tau,n}\cdots S^I_{\tau,1}w(0)-w(n\tau)\mid(\varphi,\psi)\big)_{L^2}\big|\\
&\quad\le c\tau^3\sum_{k=0}^{n-1}\Big(\|w(k\tau)\|_{D\left((M^{(0)}_{\mathrm{div}})^2\right)}+\|(\mathbf{J}_0(k\tau),0)\|_{D(M^{(0)}_{\mathrm{div}})}+\|(\mathbf{J}_0'(k\tau),0)\|_X\\
&\qquad\qquad+\int_{k\tau}^{(k+1)\tau}\|(\mathbf{J}_0''(s),0)\|_X\,\mathrm{d}s\Big)e^{6\kappa_Y n\tau}\,\|(\varphi,\psi)\|_{H^1}\\
&\quad\le C\tau^2T\Big(\|w_0\|_{D\left((M^{(0)}_{\mathrm{div}})^2\right)}+\|(\mathbf{J}_0,0)\|_{C\left([0,T],D(M^{(0)}_{\mathrm{div}})\right)}\\
&\qquad\qquad+\|(\mathbf{J}_0,0)\|_{C^2\left([0,T],X^{(0)}_{\mathrm{div}}\right)}\Big)e^{6\kappa_Y T}\,\|(\varphi,\psi)\|_{H^1}.
\end{aligned}$$

Thereby, C only depends on $\|\varepsilon\|_{W^{1,\infty}}$, $\|\mu\|_{W^{1,\infty}}$, $\|\sigma\|_{W^{1,\infty}}$ and δ. □

9.3. Near preservation of the divergence conditions in H^{-1}

Due to Proposition 7.19, the solution of (7.1) fulfils the divergence conditions

$$\begin{aligned}
\operatorname{div}(\varepsilon\mathbf{E}(t))&=\operatorname{div}(\varepsilon\mathbf{E}_0)-\int_0^t\operatorname{div}(\sigma\mathbf{E}(s)+\mathbf{J}_0(s))\,\mathrm{d}s,\\
\operatorname{div}(\mu\mathbf{H}(t))&=0
\end{aligned}$$

in $H^{-1}(Q)$ for all $t\in[0,T]$ if $(\mathbf{E}_0,\mathbf{H}_0)\in D(M)$ and

$$(\mathbf{J}_0,0)\in C\big([0,T],D(M)\big)+C^1\big([0,T],X\big).$$

Theorem 9.6. *Let $T>0$, $(\boldsymbol{E}_0,\boldsymbol{H}_0)\in D(B_Y)$ and $(\boldsymbol{J}_0,0)\in C\big([0,T],D(A_Y)\big)\cap C^1\big([0,T],X\big)$. Then there exists a bound $\tau_0\in(0,T]$ on the time step size such that the numerical solution fulfils the divergence conditions in $H^{-1}(Q)$ for time step sizes $\tau\in(0,\tau_0]$ up to order one in τ; more precisely, for all $\tau\in(0,\tau_0]$ and $N\in\mathbb{N}$ with $N\tau\le T$ we have*

$$\begin{aligned}
&\Big\|\big(\operatorname{div}(\varepsilon\boldsymbol{E}_N),\operatorname{div}(\mu\boldsymbol{H}_N)\big)-\big(\operatorname{div}(\varepsilon\boldsymbol{E}_0),0\big)\\
&\quad+\sum_{k=0}^{N-1}\tfrac{\tau}{2}\big(\operatorname{div}(\tfrac{\sigma}{2}\boldsymbol{E}_{k+1}+\sigma\boldsymbol{E}_{k+1/2}+\tfrac{\sigma}{2}\boldsymbol{E}_k),0\big)+\int_0^{N\tau}\big(\operatorname{div}(\boldsymbol{J}_0(s)),0\big)\,\mathrm{d}s\Big\|_{H^{-1}}
\end{aligned}$$

$$\leq C\tau\Bigg(\int_0^T \|(\boldsymbol{J}_0'(s),0)\|_{L^2}\,\mathrm{d}s + e^{6\kappa_Y T}\Big(\big(\|(\boldsymbol{E}_0,\boldsymbol{H}_0)\|_{H^1} + \tau\,\|B_Y(\boldsymbol{E}_0,\boldsymbol{H}_0)\|_{H^1}\big) \tag{9.5}$$
$$+ T\sup_{t\in[0,T]}\big(\|(\boldsymbol{J}_0(t),0)\|_{H^1} + \tau\,\big\|A_Y\big(\tfrac{1}{\varepsilon}\boldsymbol{J}_0(t),0\big)\big\|_{H^1}\big)\Big)\Bigg)$$

for a constant $C \geq 0$ only depending on $\|\varepsilon\|_{W^{1,\infty}}$, $\|\mu\|_{W^{1,\infty}}$, $\|\sigma\|_{W^{1,\infty}}$ and δ.

Remark 9.7. *The proof of Theorem 9.6 below also shows that under the same assumptions we have for all $\tau \in (0,\tau_0]$ and $N \in \mathbb{N}$ with $N\tau \leq T$ the estimate*

$$\begin{aligned}
&\Big\|\big(\mathrm{div}(\varepsilon\boldsymbol{E}_N),\mathrm{div}(\mu\boldsymbol{H}_N)\big) - \big(\mathrm{div}(\varepsilon\boldsymbol{E}_0),0\big)\\
&\quad+\sum_{k=0}^{N-1}\tfrac{\tau}{2}\big(\mathrm{div}(\tfrac{\sigma}{2}\boldsymbol{E}_{k+1}+\sigma\boldsymbol{E}_{k+1/2}+\tfrac{\sigma}{2}\boldsymbol{E}_k),0\big)\\
&\quad+\sum_{k=0}^{N-1}\tfrac{\tau}{2}\big(\mathrm{div}(\boldsymbol{J}_0(t_k)+\boldsymbol{J}_0(t_{k+1})),0\big)\Big\|_{H^{-1}}\\
&\leq C\tau\Bigg(e^{6\kappa_Y T}\Big(\big(\|(\boldsymbol{E}_0,\boldsymbol{H}_0)\|_{H^1}+\tau\,\|B_Y(\boldsymbol{E}_0,\boldsymbol{H}_0)\|_{H^1}\big)\\
&\quad+T\sup_{t\in[0,T]}\big(\|\boldsymbol{J}_0(t)\|_{H^1}+\tau\,\|A_Y(\boldsymbol{J}_0(t),0)\|_{H^1}\big)\Big)\Bigg)
\end{aligned}\tag{9.6}$$

for a constant $C \geq 0$ only depending on $\|\varepsilon\|_{W^{1,\infty}}$, $\|\mu\|_{W^{1,\infty}}$, $\|\sigma\|_{W^{1,\infty}}$ and δ. This version of Theorem 9.6 can be used for numerical confirmations.

Proof (of Theorem 9.6):
Let $\tau_0 := \min\{\frac{1}{2\kappa_Y},\widetilde{\tau}\}$ with the bound $\widetilde{\tau}$ on the time step size from Proposition 8.7 and let $\tau \in (0,\tau_0]$. In the following we write $t_k := k\tau$ for $k \in \mathbb{N}$ and make frequently use of the assumption $\tau \leq 1$. Let $n \in \mathbb{N}$ and $w_0 := (\mathbf{E}_0,\mathbf{H}_0)$.

We first show a recursion formula for the divergence of numerical solution and then insert it into itself to obtain a closed, but nevertheless implicit, formula for the divergence. Afterwards we bring all terms that approximate the divergence condition to one side of the equation and estimate the error of the approximation of the integral as well as the other summands.

1) We have by Remark 8.13 and the identities (8.10) and (8.9) that $(\mathbf{E}_n,\mathbf{H}_n) \in D(B_Y)$ and

$$(\mathbf{E}_{n+1/2},\mathbf{H}_{n+1/2}) = \big(I-\tfrac{\tau}{2}A_Y\big)^{-1}\big(I+\tfrac{\tau}{2}B_Y\big)(\mathbf{E}_n,\mathbf{H}_n) \in D(A_Y),$$

and therefore

$$\begin{aligned}
&\big((1+\tfrac{\sigma\tau}{4\varepsilon})\mathbf{E}_{n+1/2},\mathbf{H}_{n+1/2}\big) - \tfrac{\tau}{2}\big(\tfrac{1}{\varepsilon}C_1\mathbf{H}_{n+1/2},\tfrac{1}{\mu}C_2\mathbf{E}_{n+1/2}\big)\\
&\quad=\big((1-\tfrac{\sigma\tau}{4\varepsilon})\mathbf{E}_n,\mathbf{H}_n\big) - \tfrac{\tau}{2}\big(\tfrac{1}{\varepsilon}C_2\mathbf{H}_n,\tfrac{1}{\mu}C_1\mathbf{E}_n\big)
\end{aligned}$$

in Y. Reordering the terms and plugging the equation into itself gives in $L^2(Q)^6$

$$\begin{aligned}\big((1+\tfrac{\sigma\tau}{4\varepsilon})\mathbf{E}_{n+1/2},\mathbf{H}_{n+1/2}\big) &= \tfrac{\tau}{2}\Big(\tfrac{1}{\varepsilon}C_1\big(\tfrac{\tau}{2\mu}C_2\mathbf{E}_{n+1/2}+\mathbf{H}_n-\tfrac{\tau}{2\mu}C_1\mathbf{E}_n\big),\\ &\qquad \tfrac{1}{\mu}C_2\big(\tfrac{\tau}{2\varepsilon}C_1\mathbf{H}_{n+1/2}+(1-\tfrac{\sigma\tau}{4\varepsilon})\mathbf{E}_n-\tfrac{\tau}{2\varepsilon}C_2\mathbf{H}_n\big)\Big)\\ &\quad+\tfrac{\tau}{2}\big(0,-\tfrac{1}{\mu}C_2\tfrac{\sigma\tau}{4\varepsilon}\mathbf{E}_{n+1/2}\big)\\ &\quad+\big((1-\tfrac{\sigma\tau}{4\varepsilon})\mathbf{E}_n,\mathbf{H}_n\big)-\tfrac{\tau}{2}\big(\tfrac{1}{\varepsilon}C_2\mathbf{H}_n,\tfrac{1}{\mu}C_1\mathbf{E}_n\big).\end{aligned}$$

Hence, recalling (8.11) and using $\operatorname{curl}=C_1-C_2$,

$$\begin{aligned}&\big(\varepsilon\mathbf{E}_{n+1/2}-\tfrac{\tau^2}{4}D_\mu^{(1)}\mathbf{E}_{n+1/2},\mu\mathbf{H}_{n+1/2}-\tfrac{\tau^2}{4}D_\varepsilon^{(2)}\mathbf{H}_{n+1/2}\big)\\ &\quad=\big(\varepsilon\mathbf{E}_n-\tfrac{\tau^2}{4}C_1\tfrac{1}{\mu}C_1\mathbf{E}_n,\mu\mathbf{H}_n-\tfrac{\tau^2}{4}C_2\tfrac{1}{\varepsilon}C_2\mathbf{H}_n\big)\\ &\qquad-\tfrac{\tau}{2}\big(0,C_2\tfrac{\sigma\tau}{4\varepsilon}(\mathbf{E}_{n+1/2}+\mathbf{E}_n)\big)\\ &\qquad-\big(\tfrac{\sigma\tau}{4}(\mathbf{E}_{n+1/2}+\mathbf{E}_n),0\big)+\tfrac{\tau}{2}(\operatorname{curl}\mathbf{H}_n,-\operatorname{curl}\mathbf{E}_n)\end{aligned}\tag{9.7}$$

in $L^2(Q)^6$. From

$$\begin{aligned}(\mathbf{E}_{n+1},\mathbf{H}_{n+1})=\big(I-\tfrac{\tau}{2}B_Y\big)^{-1}\big(I+\tfrac{\tau}{2}A_Y\big)\big((\mathbf{E}_{n+1/2},\mathbf{H}_{n+1/2})\\ -\tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1}),0)\big),\end{aligned}$$

see (8.10) and (8.9), we get

$$\begin{aligned}&\big((1+\tfrac{\sigma\tau}{4\varepsilon})\mathbf{E}_{n+1},\mathbf{H}_{n+1}\big)+\tfrac{\tau}{2}\big(\tfrac{1}{\varepsilon}C_2\mathbf{H}_{n+1},\tfrac{1}{\mu}C_1\mathbf{E}_{n+1}\big)\\ &\quad=\big((1-\tfrac{\sigma\tau}{4\varepsilon})\mathbf{E}_{n+1/2},\mathbf{H}_{n+1/2}\big)+\tfrac{\tau}{2}\big(\tfrac{1}{\varepsilon}C_1\mathbf{H}_{n+1/2},\tfrac{1}{\mu}C_2\mathbf{E}_{n+1/2}\big)\\ &\qquad-\big((1-\tfrac{\sigma\tau}{4\varepsilon})\tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1})),\tfrac{\tau}{2\mu}C_2\tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1}))\big)\end{aligned}$$

in Y. Again we reorder the terms and plug the equation into itself, getting

$$\begin{aligned}&\big((1+\tfrac{\sigma\tau}{4\varepsilon})\mathbf{E}_{n+1},\mathbf{H}_{n+1}\big)\\ &\quad=-\tfrac{\tau}{2}\Big(\tfrac{1}{\varepsilon}C_2\big(-\tfrac{\tau}{2\mu}C_1\mathbf{E}_{n+1}+\mathbf{H}_{n+1/2}+\tfrac{\tau}{2\mu}C_2\mathbf{E}_{n+1/2}\big),\\ &\qquad\tfrac{1}{\mu}C_1\big(-\tfrac{\tau}{2\varepsilon}C_2\mathbf{H}_{n+1}+(1-\tfrac{\sigma\tau}{4\varepsilon})\mathbf{E}_{n+1/2}+\tfrac{\tau}{2\varepsilon}C_1\mathbf{H}_{n+1/2}\big)\Big)\\ &\qquad+\big(0,\tfrac{\tau}{2\mu}C_1\tfrac{\sigma\tau}{4\varepsilon}\mathbf{E}_{n+1}\big)\\ &\qquad-\tfrac{\tau}{2}\Big(-\tfrac{1}{\varepsilon}C_2\big(\tfrac{\tau}{2\mu}C_2\tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1}))\big),\\ &\qquad\quad-\tfrac{1}{\mu}C_1\big((1-\tfrac{\sigma\tau}{4\varepsilon})\tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1}))\big)\Big)\\ &\qquad+\big((1-\tfrac{\sigma\tau}{4\varepsilon})\mathbf{E}_{n+1/2},\mathbf{H}_{n+1/2}\big)+\tfrac{\tau}{2}\big(\tfrac{1}{\varepsilon}C_1\mathbf{H}_{n+1/2},\tfrac{1}{\mu}C_2\mathbf{E}_{n+1/2}\big)\\ &\qquad-\big((1-\tfrac{\sigma\tau}{4\varepsilon})\tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1})),\tfrac{\tau}{2\mu}C_2\tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1}))\big)\end{aligned}$$

in $L^2(Q)^6$. This yields, again with (8.11) and $\operatorname{curl}=C_1-C_2$,

$$\big(\varepsilon\mathbf{E}_{n+1}-\tfrac{\tau^2}{4}D_\mu^{(2)}\mathbf{E}_{n+1},\mu\mathbf{H}_{n+1}-\tfrac{\tau^2}{4}D_\varepsilon^{(1)}\mathbf{H}_{n+1}\big)$$

$$
\begin{aligned}
&= \big(\varepsilon\mathbf{E}_{n+1/2} - \tfrac{\tau^2}{4}C_2\tfrac{1}{\mu}C_2\mathbf{E}_{n+1/2}, \mu\mathbf{H}_{n+1/2} - \tfrac{\tau^2}{4}C_1\tfrac{1}{\varepsilon}C_1\mathbf{H}_{n+1/2}\big)\\
&\quad - \big(\tfrac{\sigma\tau}{4}(\mathbf{E}_{n+1} + \mathbf{E}_{n+1/2}), 0\big) + \tfrac{\tau}{2}(\operatorname{curl}\mathbf{H}_{n+1/2}, -\operatorname{curl}\mathbf{E}_{n+1/2})\\
&\quad + \tfrac{\tau}{2}\big(0, C_1\tfrac{\sigma\tau}{4\varepsilon}(\mathbf{E}_{n+1/2} + \mathbf{E}_{n+1})\big)\\
&\quad + \tfrac{\tau}{2}\big(C_2\tfrac{\tau}{2\mu}C_2\tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1})), -C_1\tfrac{\sigma\tau}{4\varepsilon}\tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1}))\big)\\
&\quad - \tfrac{\tau}{2}\big((1 - \tfrac{\sigma\tau}{4\varepsilon})(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1})), 0\big)\\
&\quad + \big(0, \operatorname{curl}\big(\tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1}))\big)\big)
\end{aligned}
\tag{9.8}
$$

in $L^2(Q)^6$. Let $\varphi \in H_0^2(Q)$. The identity $\operatorname{curl}\nabla = 0$ yields for all $v \in H^1(Q)^3$ with $C_2 v \in H^1(Q)^3$ and the equations (7.26) and (7.28) that

$$
\begin{aligned}
\big(D_\varepsilon^{(1)} v \mid \nabla\varphi\big)_{L^2} &= \big(\tfrac{1}{\varepsilon}C_2 v, -C_2\nabla\varphi\big)_{L^2}\\
&= \big(\tfrac{1}{\varepsilon}C_2 v, (C_1 - C_2)\nabla\varphi\big)_{L^2} - \big(\tfrac{1}{\varepsilon}C_2 v, C_1\nabla\varphi\big)_{L^2}\\
&= \big(C_2\tfrac{1}{\varepsilon}C_2 u \mid \nabla\varphi\big)_{L^2}.
\end{aligned}
$$

So, using the density of $\nabla H_0^2(Q)$ in $L^2(Q)^3$ and the continuity of $\operatorname{div} : L^2(Q)^3 \to H^{-1}(Q)^3$, we have in the distributional sense

$$\operatorname{div} D_\varepsilon^{(1)} v = \operatorname{div} C_2\tfrac{1}{\varepsilon}C_2 v$$

and, shown analogously,

$$\operatorname{div} D_\mu^{(2)} u = \operatorname{div} C_1\tfrac{1}{\mu}C_1 u$$

for all $u \in H^1(Q)^3$ with $C_1 u \in H^1(Q)^3$. Together with $0 = \operatorname{div}\operatorname{curl} = \operatorname{div} C_1 - \operatorname{div} C_2$ in the distributional sense we get in $H^{-1}(Q)^6$ for $n \geq 1$ by (9.8) and (9.7) that

$$
\begin{aligned}
&\Big(\operatorname{div}\big(\varepsilon\mathbf{E}_{n+1} - \tfrac{\tau^2}{4}D_\mu^{(2)}\mathbf{E}_{n+1}\big), \operatorname{div}\big(\mu\mathbf{H}_{n+1} - \tfrac{\tau^2}{4}D_\varepsilon^{(1)}\mathbf{H}_{n+1}\big)\Big)\\
&\quad = \Big(\operatorname{div}\big(\varepsilon\mathbf{E}_{n+1/2} - \tfrac{\tau^2}{4}D_\mu^{(1)}\mathbf{E}_{n+1/2}\big), \operatorname{div}\big(\mu\mathbf{H}_{n+1/2} - \tfrac{\tau^2}{4}D_\varepsilon^{(2)}\mathbf{H}_{n+1/2}\big)\Big)\\
&\qquad - \big(\operatorname{div}(\tfrac{\sigma\tau}{4}(\mathbf{E}_{n+1} + \mathbf{E}_{n+1/2})), 0\big) + \tfrac{\tau}{2}\Big(0, \operatorname{div}\big(C_1\tfrac{\sigma\tau}{4\varepsilon}(\mathbf{E}_{n+1/2} + \mathbf{E}_{n+1})\big)\Big)\\
&\qquad + \tfrac{\tau}{2}\Big(\operatorname{div}\big(C_2\tfrac{\tau}{2\mu}C_2\tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1}))\big),\\
&\qquad\qquad - \operatorname{div}\big(C_1\tfrac{\sigma\tau^2}{8\varepsilon^2}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1}))\big)\Big)\\
&\qquad - \tfrac{\tau}{2}\Big(\operatorname{div}\big((1 - \tfrac{\sigma\tau}{4\varepsilon})(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1}))\big), 0\Big)\Big)\\
&\quad = \Big(\operatorname{div}\big(\varepsilon\mathbf{E}_n - \tfrac{\tau^2}{4}D_\mu^{(2)}\mathbf{E}_n\big), \operatorname{div}\big(\mu\mathbf{H}_n - \tfrac{\tau^2}{4}D_\varepsilon^{(1)}\mathbf{H}_n\big)\Big)\\
&\qquad - \Big(\operatorname{div}\big(\tfrac{\sigma\tau}{4}\mathbf{E}_{n+1} + \tfrac{\sigma\tau}{2}\mathbf{E}_{n+1/2} + \tfrac{\sigma\tau}{4}\mathbf{E}_n\big), 0\Big)\\
&\qquad - \tfrac{\tau}{2}\Big(\operatorname{div}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1})), 0\Big)\\
&\qquad + \tfrac{\tau}{2}\Big(0, \operatorname{div}\big(C_1\tfrac{\sigma\tau}{4\varepsilon}(\mathbf{E}_{n+1} - \mathbf{E}_n)\big)\Big)
\end{aligned}
$$

$$
\begin{aligned}
&+ \tfrac{\tau}{2}\Big(\tfrac{\tau}{2}\operatorname{div}\big(D_\mu^{(1)}\tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1}))\big),\\
&\qquad -\operatorname{div}\big(C_1\tfrac{\sigma\tau^2}{8\varepsilon^2}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1}))\big)\Big)\\
&+ \tfrac{\tau}{2}\Big(\operatorname{div}\big(\tfrac{\sigma\tau}{4\varepsilon}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1}))\big),0\Big)\Big).
\end{aligned}
$$

Thus, we get for $N \geq 1$ by inserting this relation inductively into itself

$$
\begin{aligned}
&\Big(\operatorname{div}\big(\varepsilon\mathbf{E}_N - \tfrac{\tau^2}{4}D_\mu^{(2)}\mathbf{E}_N\big),\operatorname{div}\big(\mu\mathbf{H}_N - \tfrac{\tau^2}{4}D_\varepsilon^{(1)}\mathbf{H}_N\big)\Big)\\
&\quad=\Big(\operatorname{div}\big(\varepsilon\mathbf{E}_0 - \tfrac{\tau^2}{4}D_\mu^{(2)}\mathbf{E}_0\big),\operatorname{div}\big(\mu\mathbf{H}_0 - \tfrac{\tau^2}{4}D_\varepsilon^{(1)}\mathbf{H}_0\big)\Big)\\
&\qquad-\sum_{n=0}^{N-1}\Big(\operatorname{div}\big(\tfrac{\sigma\tau}{4}\mathbf{E}_{n+1}+\tfrac{\sigma\tau}{2}\mathbf{E}_{n+1/2}+\tfrac{\sigma\tau}{4}\mathbf{E}_n\big),0\Big)\\
&\qquad-\sum_{n=0}^{N-1}\tfrac{\tau}{2}\big(\operatorname{div}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1})),0\big)\\
&\qquad+\sum_{n=0}^{N-1}\tfrac{\tau^2}{8}\Big(0,\operatorname{div}\big(C_1\tfrac{\sigma}{\varepsilon}(\mathbf{E}_{n+1}-\mathbf{E}_n)\big)\Big)\\
&\qquad+\sum_{n=0}^{N-1}\tfrac{\tau^2}{4}\Big(\operatorname{div}\big(D_\mu^{(1)}\tfrac{\tau}{2\varepsilon}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1}))\big),\\
&\qquad\qquad-\operatorname{div}\big(C_1\tfrac{\sigma\tau}{4\varepsilon^2}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1}))\big)\Big)\\
&\qquad+\sum_{n=0}^{N-1}\tfrac{\tau}{2}\Big(\operatorname{div}\big(\tfrac{\sigma\tau}{4\varepsilon}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1}))\big),0\Big)
\end{aligned}
$$

in $H^{-1}(Q)^6$. Reordering these terms yields with $\operatorname{div}(\mu\mathbf{H}_0)=0$ that

$$
\begin{aligned}
&\big(\operatorname{div}(\varepsilon\mathbf{E}_N),\operatorname{div}(\mu\mathbf{H}_N)\big)-\big(\operatorname{div}(\varepsilon\mathbf{E}_0),0\big)\\
&\quad+\sum_{n=0}^{N-1}\big(\operatorname{div}(\tfrac{\sigma\tau}{4}\mathbf{E}_{n+1}+\tfrac{\sigma\tau}{2}\mathbf{E}_{n+1/2}+\tfrac{\sigma\tau}{4}\mathbf{E}_n),0\big)\\
&\quad+\sum_{n=0}^{N-1}\tfrac{\tau}{2}\big(\operatorname{div}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1})),0\big)\\
&=\Big(\operatorname{div}\big(\tfrac{\tau^2}{4}D_\mu^{(2)}\mathbf{E}_N\big),\operatorname{div}\big(\tfrac{\tau^2}{4}D_\varepsilon^{(1)}\mathbf{H}_N\big)\Big)-\Big(\operatorname{div}\big(\tfrac{\tau^2}{4}D_\mu^{(2)}\mathbf{E}_0\big),\operatorname{div}\big(\tfrac{\tau^2}{4}D_\varepsilon^{(1)}\mathbf{H}_0\big)\Big)\\
&\quad+\tfrac{\tau^2}{8}\Big(0,\operatorname{div}\big(C_1\tfrac{\sigma}{\varepsilon}(\mathbf{E}_N-\mathbf{E}_0)\big)\Big)\\
&\quad+\sum_{n=0}^{N-1}\tfrac{\tau^3}{16}\Big(\operatorname{div}\big(D_\mu^{(1)}\tfrac{1}{\varepsilon}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1}))\big),\\
&\qquad\qquad-\operatorname{div}\big(C_1\tfrac{\sigma}{\varepsilon^2}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1}))\big)\Big)\\
&\quad+\sum_{n=0}^{N-1}\tfrac{\tau^2}{8}\Big(\operatorname{div}\big(\tfrac{\sigma}{\varepsilon}(\mathbf{J}_0(t_n)+\mathbf{J}_0(t_{n+1}))\big),0\Big)
\end{aligned}
\tag{9.9}
$$

in $H^{-1}(Q)^6$.

2) We observe that the absolute value of the left-hand side of (9.9) is the left-hand side of (9.5) with the integral replaced by the trapezoidal quadrature rule. With the trapezoidal rule and the assumption $(\mathbf{J}_0, 0) \in C^1\big([0,\infty), L^2(Q)^6\big)$ we have

$$\begin{aligned}
&\Big\| \Big(\sum_{n=0}^{N-1} \tfrac{\tau}{2}\big(\operatorname{div}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1})\big) - \int_0^{N\tau} \operatorname{div}(\mathbf{J}_0(s))\,\mathrm{d}s, 0 \Big) \Big\|_{H^{-1}} \\
&\quad \le c \Big\| \sum_{n=0}^{N-1} \Big(\tfrac{\tau}{2}\big(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1})\big) - \int_{t_n}^{t_{n+1}} \mathbf{J}_0(s)\,\mathrm{d}s \Big) \Big\|_{L^2} \\
&\quad \le c \Big\| \sum_{n=0}^{N-1} \Big(\int_{t_n}^{\frac{1}{2}(t_n+t_{n+1})} (\mathbf{J}_0(t_n) - \mathbf{J}_0(s))\,\mathrm{d}s + \int_{\frac{1}{2}(t_n+t_{n+1})}^{t_{n+1}} (\mathbf{J}_0(t_{n+1}) - \mathbf{J}_0(s))\,\mathrm{d}s \Big) \Big\|_{L^2} \\
&\quad \le c \sum_{n=0}^{N-1} \tau \int_{t_n}^{t_{n+1}} \|\mathbf{J}_0'(s)\|_{L^2}\,\mathrm{d}s \le c\tau \int_0^T \|\mathbf{J}_0'(s)\|_{L^2}\,\mathrm{d}s.
\end{aligned}$$

Thus, it remains to bound the right-hand side of (9.9).

3) For $n \ge 1$ we have by (8.10) in Y the formulation

$$\begin{aligned}
(\mathbf{E}_n, \mathbf{H}_n) &= S^I_{\tau,n} \cdots S^I_{\tau,1} (\mathbf{E}_0, \mathbf{H}_0) \\
&= \big(I - \tfrac{\tau}{2} B_Y\big)^{-1} \gamma_{\tau/2}(A_Y) \big(\gamma_{\tau/2}(B_Y)\gamma_{\tau/2}(A_Y)\big)^{n-1} \big(I + \tfrac{\tau}{2} B_Y\big)(\mathbf{E}_0, \mathbf{H}_0) \\
&\quad - \sum_{k=0}^{n-1} \big(I - \tfrac{\tau}{2} B_Y\big)^{-1} \big(\gamma_{\tau/2}(A_Y)\gamma_{\tau/2}(B_Y)\big)^k \big(I + \tfrac{\tau}{2} A_Y\big) \cdot \\
&\qquad \cdot \tfrac{\tau}{2\varepsilon} \big(\mathbf{J}_0(t_{n-k-1}) + \mathbf{J}_0(t_{n-k}), 0\big).
\end{aligned} \tag{9.10}$$

Observe that

$$\begin{aligned}
\big(D^{(2)}_\mu \mathbf{E}_N, D^{(1)}_\varepsilon \mathbf{H}_N\big) &= - \begin{pmatrix} 0 & C_2 \\ C_1 & 0 \end{pmatrix} B_0 S^I_{\tau,N} \cdots S^I_{\tau,1} (\mathbf{E}_0, \mathbf{H}_0) \\
&= \begin{pmatrix} \varepsilon I & 0 \\ 0 & \mu I \end{pmatrix} B_0^2 S^I_{\tau,N} \cdots S^I_{\tau,1} (\mathbf{E}_0 \mathbf{H}_0) \\
&= \begin{pmatrix} \varepsilon I & 0 \\ 0 & \mu I \end{pmatrix} \Big(B_Y + \begin{pmatrix} \frac{\sigma}{2\varepsilon} I & 0 \\ 0 & 0 \end{pmatrix} \Big)^2 \cdot \\
&\quad \cdot \Big(\big(I - \tfrac{\tau}{2} B_Y\big)^{-1} \gamma_{\tau/2}(A_Y) \big(\gamma_{\tau/2}(B_Y)\gamma_{\tau/2}(A)\big)^{N-1} \big(I + \tfrac{\tau}{2} B_Y\big)(\mathbf{E}_0, \mathbf{H}_0) \\
&\qquad + \sum_{k=0}^{N-1} \big(I - \tfrac{\tau}{2} B_Y\big)^{-1} \big(\gamma_{\tau/2}(A_Y)\gamma_{\tau/2}(B_Y)\big)^k \big(I + \tfrac{\tau}{2} A_Y\big) \cdot \\
&\qquad \cdot \tfrac{\tau}{2\varepsilon} \big(\mathbf{J}_0(t_{N-k-1}) + \mathbf{J}_0(t_{N-k}), 0\big) \Big)
\end{aligned}$$

in $L^2(Q)^6$. We thus deduce

$$\begin{aligned}
&\big\|\big(\operatorname{div} D_\mu^{(2)}\mathbf{E}_N, \operatorname{div} D_\varepsilon^{(1)}\mathbf{H}_N\big)\big\|_{H^{-1}} \le c\,\big\|\big(D_\mu^{(2)}\mathbf{E}_N, D_\varepsilon^{(1)}\mathbf{H}_N\big)\big\|_{L^2} \\
&\quad\le \Bigg\| \begin{pmatrix} \varepsilon I & 0 \\ 0 & \mu I \end{pmatrix} \Big(B_Y + \begin{pmatrix} \frac{\sigma}{2\varepsilon} I & 0 \\ 0 & 0 \end{pmatrix}\Big)^2 \big(I - \tfrac{\tau}{2}B_Y\big)^{-1} \gamma_{\tau/2}(A_Y)\cdot \\
&\qquad \cdot \big(\gamma_{\tau/2}(B_Y)\gamma_{\tau/2}(A_Y)\big)^{N-1}\big(I + \tfrac{\tau}{2}B_Y\big)(\mathbf{E}_0, \mathbf{H}_0)\Bigg\|_{L^2} \qquad (9.11)\\
&\qquad + \Bigg\| \begin{pmatrix} \varepsilon I & 0 \\ 0 & \mu I \end{pmatrix} \Big(B_Y + \begin{pmatrix} \frac{\sigma}{2\varepsilon} I & 0 \\ 0 & 0 \end{pmatrix}\Big)^2 (I - \tfrac{\tau}{2}B_Y)^{-1} \sum_{k=0}^{N-1} \big(\gamma_{\tau/2}(A_Y)\gamma_{\tau/2}(B_Y)\big)^k \cdot \\
&\qquad\quad \cdot (I + \tfrac{\tau}{2}A_Y)\tfrac{1}{\varepsilon}\tfrac{\tau}{2}\big(\mathbf{J}_0(t_{N-k-1}) + \mathbf{J}_0(t_{N-k}), 0\big)\Bigg\|_{L^2}.
\end{aligned}$$

With the identity

$$\tfrac{\tau}{2}B_Y\big(I - \tfrac{\tau}{2}B_Y\big)^{-1} = \big(I - \tfrac{\tau}{2}B_Y\big)^{-1} - I$$

on Y we get by Proposition 8.7 that

$$\begin{aligned}
&\Bigg\| \frac{\tau^2}{4} \begin{pmatrix} \varepsilon I & 0 \\ 0 & \mu I \end{pmatrix} B_Y^2 \big(I - \tfrac{\tau}{2}B_Y\big)^{-1} \gamma_{\tau/2}(A_Y)\big(\gamma_{\tau/2}(B_Y)\gamma_{\tau/2}(A)\big)^{N-1}\cdot \\
&\qquad\qquad \cdot \big(I + \tfrac{\tau}{2}B_Y\big)(\mathbf{E}_0, \mathbf{H}_0)\Bigg\|_{L^2} \\
&\quad\le \frac{\tau}{2} c\,\|B\|_{\mathcal{B}(Y,X)} \big\|\big(I - \tfrac{\tau}{2}B_Y\big)^{-1} - I\big\|_{\mathcal{B}(Y)} \big\|\gamma_{\tau/2}(A_Y)\big\|_{\mathcal{B}(Y)} \cdot \\
&\qquad \cdot \Big(\big\|\gamma_{\tau/2}(B_Y)\big\|_{\mathcal{B}(Y)} \big\|\gamma_{\tau/2}(A_Y)\big\|_{\mathcal{B}(Y)}\Big)^{N-1} \big\|\big(I + \tfrac{\tau}{2}B_Y\big)(\mathbf{E}_0, \mathbf{H}_0)\big\|_{H^1} \\
&\quad\le \frac{3\tau}{2} c e^{3(2N-1)\kappa_Y \tau} \big\|\big(I + \tfrac{\tau}{2}B_Y\big)(\mathbf{E}_0, \mathbf{H}_0)\big\|_{H^1} \\
&\quad\le c\tau e^{6\kappa_Y T}\big(\|w_0\|_{H^1} + \tau\,\|B_Y w_0\|_{H^1}\big),
\end{aligned}$$

where c depends on $\|\varepsilon\|_{W^{1,\infty}}$, $\|\mu\|_{W^{1,\infty}}$, $\|\sigma\|_{W^{1,\infty}}$ and δ. With similar, but easier, estimates for the other summands of $\Big(B_Y + \begin{pmatrix} \frac{\sigma}{2\varepsilon} I & 0 \\ 0 & 0 \end{pmatrix}\Big)^2$ we thus have

$$\begin{aligned}
&\Bigg\| \frac{\tau^2}{4} \begin{pmatrix} \varepsilon I & 0 \\ 0 & \mu I \end{pmatrix} \Big(B_Y + \begin{pmatrix} \frac{\sigma}{2\varepsilon} I & 0 \\ 0 & 0 \end{pmatrix}\Big)^2 \cdot \\
&\qquad \cdot \big(I - \tfrac{\tau}{2}B_Y\big)^{-1} \gamma_{\tau/2}(A_Y)\big(\gamma_{\tau/2}(B_Y)\gamma_{\tau/2}(A)\big)^{N-1}\big(I + \tfrac{\tau}{2}B_Y\big)(\mathbf{E}_0, \mathbf{H}_0)\Bigg\|_{L^2} \qquad (9.12)\\
&\quad\le c\tau e^{6\kappa_Y T}\big(\|w_0\|_{H^1} + \tau\,\|B_Y w_0\|_{H^1}\big),
\end{aligned}$$

where the constant c depends on the same quantities as before. In the same way as above we get

$$\begin{aligned}
&\Big\| \frac{\tau^2}{4} \begin{pmatrix} \varepsilon I & 0 \\ 0 & \mu I \end{pmatrix} \Big(B_Y + \begin{pmatrix} \frac{\sigma}{2\varepsilon} I & 0 \\ 0 & 0 \end{pmatrix} \Big)^2 (I - \tfrac{\tau}{2} B_Y)^{-1} \sum_{k=0}^{N-1} \big(\gamma_{\tau/2}(A_Y) \gamma_{\tau/2}(B_Y) \big)^k \cdot \\
&\qquad \cdot (I + \tfrac{\tau}{2} A_Y) \tfrac{1}{\varepsilon} \tfrac{\tau}{2} \big(\mathbf{J}_0(t_{N-k-1}) + \mathbf{J}_0(t_{N-k}), 0 \big) \Big\|_{L^2} \\
&\leq c\tau e^{6\kappa_Y T} \sum_{k=0}^{N-1} \tau \big(\|(\mathbf{J}_0(t_k), 0)\|_{H^1} + \tau \big\| A_Y \big(\tfrac{1}{\varepsilon} \mathbf{J}_0(t_k), 0 \big) \big\|_{H^1} \\
&\qquad + \|(\mathbf{J}_0(t_{k+1}), 0)\|_{H^1} + \big\| A_Y \big(\tfrac{1}{\varepsilon} \mathbf{J}_0(t_{k+1}), 0 \big) \big\|_{H^1} \big) \\
&\leq cTe^{6\kappa_Y T} \tau \sup_{t \in [0,T]} \big(\|(\mathbf{J}_0(t), 0)\|_{H^1} + \tau \big\| A_Y \big(\tfrac{1}{\varepsilon} \mathbf{J}_0(t), 0 \big) \big\|_{H^1} \big),
\end{aligned} \tag{9.13}$$

with c again depending on $\|\varepsilon\|_{W^{1,\infty}}$, $\|\mu\|_{W^{1,\infty}}$, $\|\sigma\|_{W^{1,\infty}}$ and δ. Altogether we have with (9.11), (9.12) and (9.13) that

$$\begin{aligned}
&\Big\| \frac{\tau^2}{4} \big(\operatorname{div} D_\mu^{(2)} \mathbf{E}_N, \operatorname{div} D_\varepsilon^{(1)} \mathbf{H}_N \big) \Big\|_{L^2} \\
&\quad \leq ce^{6\kappa_Y T} \tau \Big(\big(\|w_0\|_{H^1} + \tau \|B_Y w_0\|_{H^1} \big) + T \sup_{t \in [0,T]} \big(\|\mathbf{J}_0(t)\|_{H^1} + \tau \big\| A_Y \big(\tfrac{1}{\varepsilon} \mathbf{J}_0(t), 0 \big) \big\|_{H^1} \big) \Big)
\end{aligned}$$

with c depending only on $\|\varepsilon\|_{W^{1,\infty}}$, $\|\mu\|_{W^{1,\infty}}$, $\|\sigma\|_{W^{1,\infty}}$ and δ.

The identity

$$\big(D_\mu^{(2)} \mathbf{E}_0, D_\varepsilon^{(1)} \mathbf{H}_0 \big) = - \begin{pmatrix} \varepsilon I & 0 \\ 0 & \mu I \end{pmatrix} B_0^2 (\mathbf{E}_0 \mathbf{H}_0)$$

in $L^2(Q)^6$ due to $B_0(\mathbf{E}_0, \mathbf{H}_0) \in Y$ gives

$$\Big\| \frac{\tau^2}{4} \big(\operatorname{div} D_\mu^{(2)} \mathbf{E}_0, \operatorname{div} D_\varepsilon^{(1)} \mathbf{H}_0 \big) \Big\|_{H^{-1}} \leq c\tau^2 \|B_0 w_0\|_{H^1} \leq c\tau^2 \big(\|B_Y w_0\|_{H^1} + \|w_0\|_{H^1} \big)$$

with c depending only on $\|\varepsilon\|_{W^{1,\infty}}$, $\|\mu\|_{W^{1,\infty}}$, $\|\sigma\|_{W^{1,\infty}}$ and δ.

4) We now estimate the remaining terms. From (9.10) we conclude with Proposition 8.7 in the same way as above that

$$\begin{aligned}
\|\mathbf{E}_n\|_{H^1} &\leq c \big\| (I - \tfrac{\tau}{2} B_Y)^{-1} \big\|_{\mathcal{B}(Y)} \big\| \gamma_{\tau/2}(A_Y) \big\|_{\mathcal{B}(Y)} \big(\big\| \gamma_{\tau/2}(B_Y) \big\|_{\mathcal{B}(Y)} \big\| \gamma_{\tau/2}(A_Y) \big\|_{\mathcal{B}(Y)} \big)^{n-1} \cdot \\
&\qquad \cdot \big\| (I + \tfrac{\tau}{2} B_Y)(\mathbf{E}_0, \mathbf{H}_0) \big\|_{H^1} \\
&\quad + c\tau \sum_{k=0}^{n-1} \big\| (I - \tfrac{\tau}{2} B_Y)^{-1} \big\|_{\mathcal{B}(Y)} \Big(\big\| \gamma_{\tau/2}(A_Y) \big\|_{\mathcal{B}(Y)} \big\| \gamma_{\tau/2}(B_Y) \big\|_{\mathcal{B}(Y)} \Big)^k \cdot \\
&\qquad\qquad \cdot \big\| (I + A_Y) \big(\tfrac{1}{\varepsilon} (\mathbf{J}_0(t_{n-k}) + \mathbf{J}_0(t_{n-k-1})), 0 \big) \big\|_{H^1} \\
&\leq ce^{6\kappa_Y T} \Big(\big(\|w_0\|_{H^1} + \tau \|B_Y w_0\|_{H^1} \big)
\end{aligned}$$

$$+ T \sup_{t\in[0,T]} \big(\|\mathbf{J}_0(t)\|_{H^1} + \tau \big\| A_Y\big(\tfrac{1}{\varepsilon}\mathbf{J}_0(t),0\big)\big\|_{H^1}\big)\Big)$$

for all $n \in \{1,\dots,N\}$, so that

$$\begin{aligned}
&\Big\| \Big(0, \tfrac{\tau^2}{8} \operatorname{div}\big(C_1 \tfrac{\sigma}{\varepsilon}(\mathbf{E}_N - \mathbf{E}_0)\big)\Big)\Big\|_{H^{-1}} \\
&\quad \le c\tau^2 \big(\|\mathbf{E}_N\|_{H^1} + \|\mathbf{E}_0\|_{H^1}\big) \\
&\quad \le c\tau^2 e^{6\kappa_Y T}\Big(\big(\|w_0\|_{H^1} + \tau \|B_Y w_0\|_{H^1}\big) + T \sup_{t\in[0,T]} \big(\|\mathbf{J}_0(t)\|_{H^1} + \tau \big\| A_Y\big(\tfrac{1}{\varepsilon}\mathbf{J}_0(t),0\big)\big\|_{H^1}\big)\Big),
\end{aligned}$$

with c again depending only on $\|\varepsilon\|_{L^\infty}$, $\|\mu\|_{L^\infty}$, $\|\sigma\|_{L^\infty}$ and δ. Furthermore, we have, using the same techniques as above,

$$\begin{aligned}
&\Big\| \tfrac{\tau^3}{16} \sum_{n=0}^{N-1} \Big(\tau \operatorname{div}\big(D_\mu^{(1)} \tfrac{2}{\varepsilon}(\mathbf{J}_0(t_n) + \mathbf{J}(t_{n+1}))\big), 0\Big)\Big\|_{H^{-1}} \\
&\qquad \le cT\tau^2 \sup_{t\in[0,T]} \big\| A_Y\big(\tfrac{1}{\varepsilon}\mathbf{J}_0(t),0\big)\big\|_{H^1},
\end{aligned}$$

$$\Big\| \tfrac{\tau^3}{16} \sum_{n=0}^{N-1} \big(0, \operatorname{div}\big(C_1 \tfrac{\sigma}{\varepsilon^2}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1}))\big)\Big)\Big\|_{H^{-1}} \le cT\tau^2 \sup_{t\in[0,T]} \|(\mathbf{J}_0(t),0)\|_{H^1}$$

and

$$\Big\| \tfrac{\tau^2}{8} \sum_{n=0}^{N-1} \Big(\operatorname{div}\big(\tfrac{\sigma}{\varepsilon}(\mathbf{J}_0(t_n) + \mathbf{J}_0(t_{n+1}))\big), 0\Big)\Big\|_{H^{-1}} \le cT\tau \sup_{t\in[0,T]} \|(\mathbf{J}_0(t),0)\|_{H^1},$$

with c each time depending only on $\|\varepsilon\|_{W^{1,\infty}}$, $\|\mu\|_{W^{1,\infty}}$, $\|\sigma\|_{W^{1,\infty}}$ and δ. □

Remark 9.8. *As mentioned in Section 1.2, the above proof shows that the quadrature rule used in* (9.5) *and* (9.6) *for* $\operatorname{div}(\sigma \boldsymbol{E}(t))$ *cannot be replaced by the Simpson rule since the weights come out of the proof.*

9.4. Near preservation of the divergence conditions in L^2

Theorem 9.9. *Let* $T > 0$, $\varepsilon, \sigma \in W^{2,3}(Q)$, $\mu \in C^2(\overline{Q})$ *and* $\partial_\nu \varepsilon = \partial_\nu \mu = \partial_\nu \sigma = 0$ *on* Γ. *Let* $(\boldsymbol{E}_0, \boldsymbol{H}_0) \in D(B_Z)$ *and* $\big(\frac{1}{\varepsilon}\boldsymbol{J}_0, 0\big) \in C\big([0,T], D(A_Z)\big) \cap C^1\big([0,T], Y\big)$. *Then there exists a* $\tau_0 \in (0,T]$ *such that the numerical solution fulfils the divergence conditions in* $L^2(Q)$ *for time step sizes* $\tau \in (0,\tau_0]$ *up to order one in* τ*; more precisely, for all* $\tau \in (0,\tau_0]$ *and* $N \in \mathbb{N}$ *with* $N\tau \le T$ *we have*

$$\Big\| \big(\operatorname{div}(\varepsilon \boldsymbol{E}_N), \operatorname{div}(\mu \boldsymbol{H}_N)\big) - \big(\operatorname{div}(\varepsilon \boldsymbol{E}_0), 0\big)$$

$$
\begin{aligned}
&+ \sum_{k=0}^{N-1} \tfrac{\tau}{2}\big(\operatorname{div}(\tfrac{\sigma}{2}\boldsymbol{E}_{k+1} + \sigma\boldsymbol{E}_{k+1/2} + \tfrac{\sigma}{2}\boldsymbol{E}_k), 0\big) + \int_0^{N\tau} \big(\operatorname{div}(\boldsymbol{J}_0(s)), 0\big)\,\mathrm{d}s\Big\|_{L^2} \\
&\leq C\tau\Bigg(\int_0^T \|(\boldsymbol{J}_0'(s), 0)\|_{H^1}\,\mathrm{d}s + e^{6\kappa_Z T}\Big(\big(\|(\boldsymbol{E}_0, \boldsymbol{H}_0)\|_{H^2} + \tau\,\|B_Z(\boldsymbol{E}_0, \boldsymbol{H}_0)\|_{H^2}\big) \\
&\qquad + T \sup_{t\in[0,T]} \big(\big\|\big(\tfrac{1}{\varepsilon}\boldsymbol{J}_0(t), 0\big)\big\|_{H^2} + \tau\,\big\|A_Z\big(\tfrac{1}{\varepsilon}\boldsymbol{J}_0(t), 0\big)\big\|_{H^2}\big)\Big)\Bigg)
\end{aligned}
$$

with a constant $C \geq 0$ only depending on $\|\varepsilon\|_{W^{1,\infty}\cap W^{2,3}}$, $\|\mu\|_{W^{1,\infty}\cap W^{2,3}}$, $\|\sigma\|_{W^{1,\infty}\cap W^{2,3}}$ and δ.

Proof:
The algebraic reformulations of the proof of Theorem 9.6 can under the assumptions of Theorem 9.9 be done with the identities being in Z instead of Y and in Y instead of $L^2(Q)^6$. We arrive at the identity (9.9). With the replacements A_Y by A_Z, B_Y by B_Z and κ_Y by κ_Z the rest of the proof is done analogously to the one of Theorem 9.6, using Proposition 8.12. □

Remark 9.10. *The proof of Theorem 9.9 also shows that under the same assumptions we have for all $\tau \in (0, \tau_0]$ and $N \in \mathbb{N}$ with $N\tau \leq T$ the estimate*

$$
\begin{aligned}
&\Big\|\big(\operatorname{div}(\varepsilon\boldsymbol{E}_N), \operatorname{div}(\mu\boldsymbol{H}_N)\big) - \big(\operatorname{div}(\varepsilon\boldsymbol{E}_0), 0\big) \\
&\qquad + \sum_{k=0}^{N-1} \tfrac{\tau}{2}\big(\operatorname{div}(\tfrac{\sigma}{2}\boldsymbol{E}_{k+1} + \sigma\boldsymbol{E}_{k+1/2} + \tfrac{\sigma}{2}\boldsymbol{E}_k), 0\big) + \sum_{k=0}^{N-1} \tfrac{\tau}{2}\big(\operatorname{div}(\boldsymbol{J}_0(t_k) + \boldsymbol{J}_0(t_{k+1})), 0\big)\Big\|_{L^2} \\
&\leq C\tau\Bigg((1+\tau)e^{6\kappa_Z T}\Big(\big(\|(\boldsymbol{E}_0, \boldsymbol{H}_0)\|_{H^2} + \tau\,\|B_Z(\boldsymbol{E}_0, \boldsymbol{H}_0)\|_{H^2}\big) \\
&\qquad + T \sup_{t\in[0,T]} \big(\|(\boldsymbol{J}_0(t), 0)\|_{H^2} + \tau\,\|A_Z(\boldsymbol{J}_0(t), 0)\|_{H^2}\big)\Big)\Bigg)
\end{aligned}
$$

with a constant $C \geq 0$ only depending on $\|\varepsilon\|_{W^{1,\infty}\cap W^{2,3}}$, $\|\mu\|_{W^{1,\infty}\cap W^{2,3}}$, $\|\sigma\|_{W^{1,\infty}\cap W^{2,3}}$ and δ. We use this version of Theorem 9.9 in Section 10.2 for numerical confirmations, see Chapter 10.

10. Numerical experiments with the ADI scheme for the Maxwell equations

In this chapter we conduct numerical experiments to confirm some of our theoretical results of Chapter 9. We give in Section 10.1 an overview over the experiments and the setting we use for them. In Section 10.2 we first deal with the situation of no electrical conductivity ($\sigma = 0$) and no external currents ($\mathbf{J}_0 = 0$). We are able to confirm the results of Section 4.4 in [37] and furthermore see that the error of the divergence is very small. Afterwards, we include conductivity ($\sigma \neq 0$) and external currents ($\mathbf{J}_0 \neq 0$). We confirm the second order convergence of the ADI scheme from Theorem 9.3, and the preservation of first order of the divergence conditions from Theorem 9.9. The experiment in Section 10.3 shows that the requirement for initial functions to be in $D(M_{\mathrm{div}}^{(2)})$ cannot be weakened to $X_{\mathrm{div}}^{(2)}$.

10.1. An overview over the numerical experiments

We do the numerical computations on the three-dimensional unit cube $Q := (0,1)^3$. We discretize it by the *Yee grid*, see [72], which is a staggered grid. The idea is that the electric and the magnetic field are evaluated on different grids. This allows an efficient implementation of the space derivatives with finite differences with a step size of half the mesh width. It does not matter for our purposes that the divergence is not discretized on points of the Yee grid. Moreover, the zero tangential trace of the electric field on Γ is incorporated into the discretization of the operators.

More precisely, we choose a maximal number N of grid points in each direction and define $y_k := k/N$ for $k = 0, \ldots, N$ and $y_{k+1/2} := (k + 1/2)/N$ for $k = 0, \ldots, N-1$. The first component of the electric field then has values on the grid points $(y_{k+1/2}, y_l, y_m)$ for $k = 1, \ldots, N-1$ and $m, l = 0, \ldots, N$. The first component of the magnetic field is defined on the points $(y_k, y_{l+1/2}, y_{m+1/2})$ for $k = 0, \ldots, N$ and $m, l = 0, \ldots, N-1$. The other components of the electric and the magnetic field are discretized on the analogous space grids. Figure 10.1 sketches the discretization.

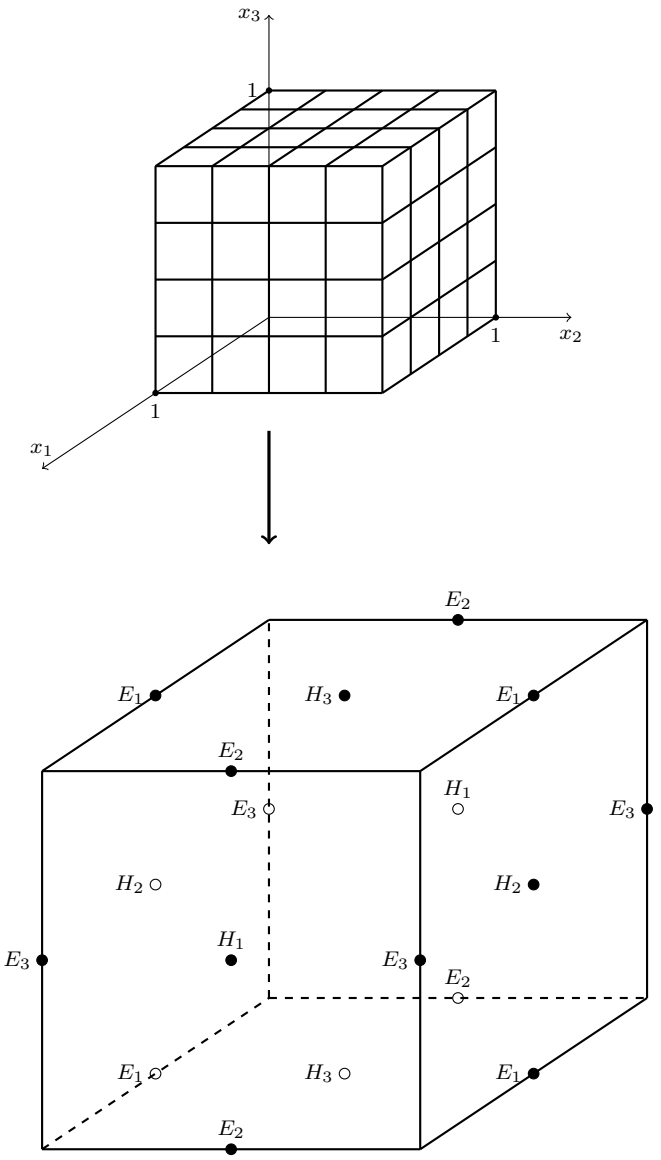


Figure 10.1.: Sketch of the partitioning of Q into the cells of the Yee grid and drawing of one cell of the Yee grid.

The time domain for our computations is the interval $[0, 1]$, which we discretize by uniform time steps. We use time step sizes of length $1/10$ for $1/640$, slightly varying from experiment to experiment. The error of the ADI method is measured by calculating the discrete L^2-norm of the error term at several equidistant time points. The error term is the difference between the result of the ADI method computation and either the exact solution (if available by an explicit formula) or a reference solution. The final error is defined as the maximum over those discrete L^2-norms. For the errors of the divergence preservation we bring all summands in the fully discretized version of (7.2) to one side.

This reads

$$\begin{aligned}\operatorname{div}(\varepsilon \mathbf{E}_N^h) - \operatorname{div}(\varepsilon \mathbf{E}_0^h) + \sum_{k=0}^{N-1} \tfrac{\tau}{2} \operatorname{div}(\tfrac{\sigma}{2}\mathbf{E}_{k+1}^h + \sigma \mathbf{E}_{k+1/2}^h + \tfrac{\sigma}{2}\mathbf{E}_k^h) \\ + \sum_{k=0}^{N-1} \tfrac{\tau}{2} \operatorname{div}(\mathbf{J}_0^h(t_k) + \mathbf{J}_0^h(t_{k+1})) = 0, \\ \operatorname{div}(\mu \mathbf{H}_N^h) = 0,\end{aligned}$$

where $\mathbf{E}_N^h$ is the (spatially discretized) result of the ADI splitting of the electric field after N time steps, and so on. We compute at several equidistant time points the discrete L^2-norm and take the maximum over those values. In Subsection 10.2.1 and for the experiment in Subsection 10.2.2 on the divergence conditions we use five time steps. For the experiments on the convergence order in the Subsection 10.2.2 and the one in Section 10.3 we use ten time steps.

10.2. Verification of the theoretical results

10.2.1. Experiments without conductivity and external current

First we treat the case that we have no conductivity ($\sigma = 0$) and no external currents ($\mathbf{J}_0 = 0$), see Section 4.4 in [37]. The parameter functions are chosen to be $\varepsilon = 1$ and $\mu = 1$. In this situation we have solutions to (7.1) that are given by explicit formulas, namely

$$u^{(1)}(t,x) = \begin{pmatrix} \sin(\pi x_2)\sin(\pi x_3)\cos(\sqrt{2}\pi t) \\ 0 \\ 0 \\ 0 \\ -\frac{1}{2}\sqrt{2}\sin(\pi x_2)\cos(\pi x_3)\sin(\sqrt{2}\pi t) \\ \frac{1}{2}\sqrt{2}\cos(\pi x_2)\sin(\pi x_3)\sin(\sqrt{2}\pi t) \end{pmatrix},$$

$$u^{(2)}(t,x) = \begin{pmatrix} 0 \\ \sin(\pi x_1)\sin(\pi x_3)\cos(\sqrt{2}\pi t) \\ 0 \\ \frac{1}{2}\sqrt{2}\sin(\pi x_1)\cos(\pi x_3)\sin(\sqrt{2}\pi t) \\ 0 \\ -\frac{1}{2}\sqrt{2}\cos(\pi x_1)\sin(\pi x_3)\sin(\sqrt{2}\pi t) \end{pmatrix}$$

and

$$u^{(3)}(t,x) = \begin{pmatrix} 0 \\ 0 \\ \sin(\pi x_1)\sin(\pi x_2)\cos(\sqrt{2}\pi t) \\ -\frac{1}{2}\sqrt{2}\sin(\pi x_1)\cos(\pi x_2)\sin(\sqrt{2}\pi t) \\ \frac{1}{2}\sqrt{2}\cos(\pi x_1)\sin(\pi x_2)\sin(\sqrt{2}\pi t) \\ 0 \end{pmatrix}.$$

The corresponding initial basis functions are

$$u_0^{(1)}(x) = u^{(1)}(0,x) = \begin{pmatrix} \sin(\pi x_2)\sin(\pi x_3) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

$$u_0^{(2)}(x) = u^{(2)}(0,x) = \begin{pmatrix} 0 \\ \sin(\pi x_1)\sin(\pi x_3) \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

and

$$u_0^{(3)}(x) = u^{(3)}(0,x) = \begin{pmatrix} 0 \\ 0 \\ \sin(\pi x_1)\sin(\pi x_2) \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

We define the initial function

$$u_0 = 2u_0^{(1)} + 3u_0^{(2)} + 4u_0^{(3)}.$$

In contrast to the authors of [37] we choose slightly different coefficients and we do not normalize the initial function in the L^2-norm. Applying M to $u_0^{(1)}$ first gives

$$Mu_0^{(1)}(x) = -u_0^{(1)}(x) + \pi \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \sin(\pi x_2)\cos(\pi x_3) \\ -\cos(\pi x_2)\sin(\pi x_3) \end{pmatrix}$$

and then

$$M^2 u_0^{(1)}(x) = -Mu_0^{(1)}(x) + \pi^2 \begin{pmatrix} 2\sin(\pi x_2)\sin(\pi x_3) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Repeating this inductively it follows that $M^m u_0^{(1)} \in D(M)$ for all $m \in \mathbb{N}_0$ and that all the trace and divergence conditions for $u_0^{(1)}$ being in $D(M_{\mathrm{div}}^{(2)})$ are satisfied. Arguing with $u_0^{(2)}$ and $u_0^{(3)}$ in the same way, we obtain $u_0 \in D(M_{\mathrm{div}}^{(2)})$. Due to the linearity of the Maxwell equations (7.1) we have that the exact solution to this initial function is

$$u(t) = 2u^{(1)}(t) + 3u^{(2)}(t) + 4u^{(3)}(t).$$

The errors between the computed approximate solutions and the exact solutions are displayed in Figure 10.2. For the larger time step sizes we see convergence of order two. For small time step sizes the spatial error dominates the total error, and the plateaus being visible indicate the space discretization errors.

In Figure 10.3 one sees the L^2-deviation of the divergence terms from zero, which is in the order of the machine accuracy. This shows that errors of the divergence preservation of numerical solutions appearing later in the experiments are caused by the errors of the numerical solution and not by the numerical method that computes the error of the divergence preservation.

10.2.2. Experiments with conductivity and external current

In this subsection we extend the setting of Subsection 10.2.1 by adding a conductivity and external currents. We conduct experiments to confirm the statements of Theorem 9.3 and 9.9, which predict a temporal convergence in L^2 of order two, and a preservation of the divergence conditions of order one in the time step size.

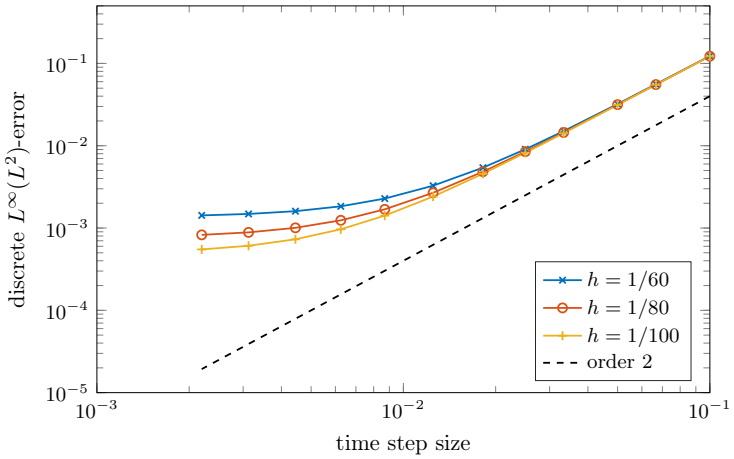


Figure 10.2.: Errors of the ADI splitting method without conductivity and external current, using a space mesh width of h.

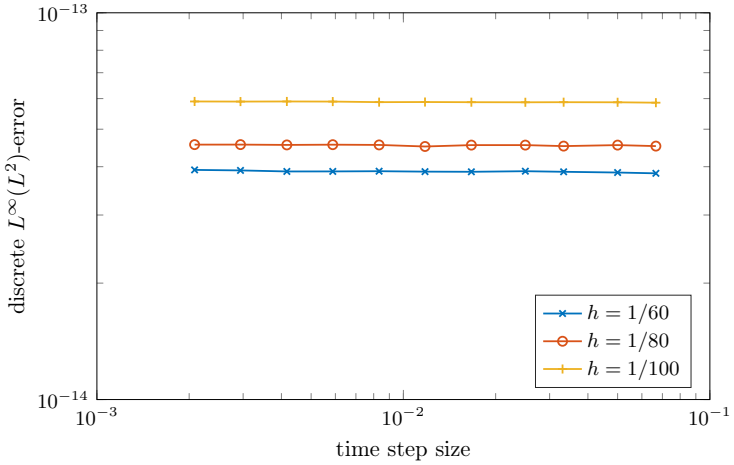


Figure 10.3.: Errors of the preservation of the divergence conditions for the case of no conductivity and external current, using a space mesh width of h.

In contrast to Subsection 10.2.1 we do not have a formula for the exact solution. Therefore, we have to compute a reference solution. We do this with the ADI scheme with the very small time step size of 1/1920.

As in Subsection 10.2.1, we use the initial function

$$(\mathbf{E}_0, \mathbf{H}_0)(x) := \begin{pmatrix} 2\sin(\pi x_2)\sin(\pi x_3) \\ 3\sin(\pi x_1)\sin(\pi x_3) \\ 4\sin(\pi x_1)\sin(\pi x_2) \\ 0 \\ 0 \\ 0 \end{pmatrix} \in D(M_{\mathrm{div}}^{(2)}) \cap D(B_Z).$$

The electric permittivity, the magnetic permeability and the conductivity density are

chosen to be

$$\varepsilon(x) := \mu(x) := 1 + g_1(x_1)g_1(x_2)g_1(x_3)$$

and

$$\sigma(x) := g_1(x_1)g_1(x_2)g_1(x_3)$$

with

$$g_1(y) := -2y^3 + 3y^2.$$

Due to $g_1'(0) = g_1'(1) = 0$ and $g_1 \geq 0$ on $[0,1]$, the functions ε, μ and σ satisfy the normal trace conditions $\partial_\nu \varepsilon = \partial_\nu \mu = \partial_\nu \sigma = 0$ on Γ and their positivity assumptions. Moreover, we have $\varepsilon, \mu, \sigma \in W^{1,\infty}\big(Q\big) \cap W^{2,3}\big(Q\big)$. As current density we use

$$\mathbf{J}_0(t,x) := \begin{pmatrix} g_2(x_2)g_2(x_3)\sin(t) \\ g_2(x_1)g_2(x_3)\cos(2t) \\ g_2(x_1)g_2(x_2)\sin(t)\cos(3t) \end{pmatrix}$$

with

$$g_2(y) := 50y^3(1-y)^3.$$

The smoothness of g_2 and all its derivatives, and the zero boundary condition of g_2, g_2' and g_2'' at $y = 0$ and $y = 1$ ensure that

$$\big(\tfrac{1}{\varepsilon}\mathbf{J}_0, 0\big) \in C^1\big([0,1], X_{\mathrm{div}}^{(2)}\big) \cap C^2\big([0,1], D(M_{\mathrm{div}}^{(0)})\big)$$

and

$$\big(\tfrac{1}{\varepsilon}\mathbf{J}_0, 0\big) \in C\big([0,1], D(A_Z)\big) \cap C^1\big([0,1], Y\big).$$

So, the requirements of Theorem 9.3 and 9.9 on the initial function, the coefficient functions and the current density are satisfied.

We first deal with the convergence order of the ADI scheme in the L^2-setting. The results, displayed in Figure 10.4, show convergence order two in the time step size as predicted by Theorem 9.3, independent of the space mesh width. In contrast to Figure 10.2 no plateaus showing the space discretization error are visible. The reason is that the solution and the reference solution are computed on the same spatial grid so that in Figure 10.4 only the time discretization error is visible.

We now compute the numerical error of the divergence conditions in the L^2-setting. The results, displayed in Figure 10.5, show (for small time step sizes) a preservation of the divergence conditions of order one in the time step size, in perfect coincidence with Theorem 9.9, independent of the space mesh width.

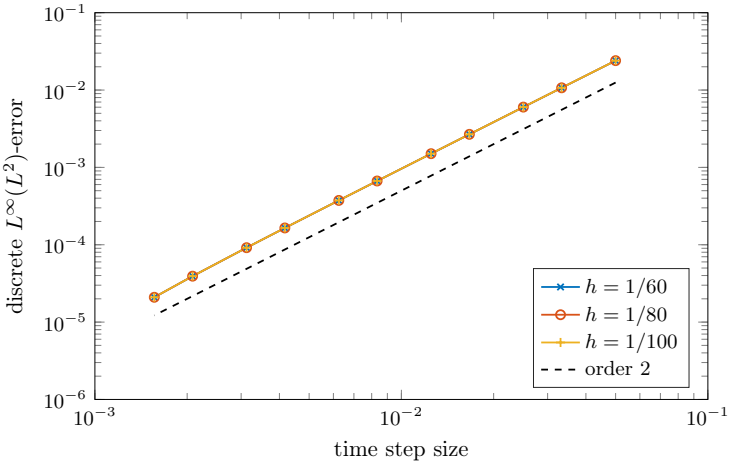


Figure 10.4.: Errors of the ADI method, using the space mesh width h.

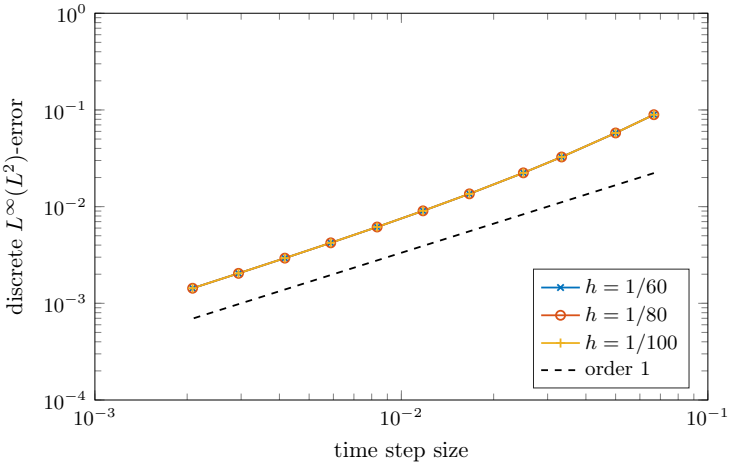


Figure 10.5.: Errors of the preservation of the divergence conditions, using the space mesh width h.

10.3. An order reduction for an initial function with low regularity

The Cauchy problem (7.34) has a unique solution in $X^{(2)}_{\mathrm{div}}$ if the initial function is in $D(M^{(2)}_{\mathrm{div}})$, see Proposition 7.21. We add an experiment which shows that the requirement that $(\mathbf{E}_0, \mathbf{H}_0)$ belongs to $D(M^3)$ that is contained in the assumption $(\mathbf{E}_0, \mathbf{H}_0) \in D(M^{(2)}_{\mathrm{div}})$ is necessary for the full convergence order in Theorem 9.3, and that $(\mathbf{E}_0, \mathbf{H}_0) \in X^{(2)}_{\mathrm{div}}$ with the included assumption $(\mathbf{E}_0, \mathbf{H}_0) \in D(M^2)$ is not sufficient.

Choose $\mathbf{J}_0$ as in the experiments in Subsection 10.2.2. Let $\sigma = \varepsilon = \mu = 1$ and

$$(\mathbf{E}_0, \mathbf{H}_0) = \begin{pmatrix} \sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3) \\ \sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3) \\ \sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3) \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

We immediately see $(\mathbf{E}_0, \mathbf{H}_0) \in D(M)$, $\operatorname{div}(\mathbf{H}_0) = 0$, $\operatorname{tr}_n(\mathbf{H}_0) = 0$ and $\operatorname{div}(\mathbf{E}_0) \in H^1(Q)$. From

$$\begin{aligned} \operatorname{div}(\mathbf{E}_0) = \pi\big(&\cos(\pi x_1)\sin(\pi x_2)\sin(\pi x_3) + \sin(\pi x_1)\cos(\pi x_2)\sin(\pi x_3) \\ &+ \sin(\pi x_1)\sin(\pi x_2)\cos(\pi x_3)\big) \end{aligned}$$

we see that $\operatorname{div}(\mathbf{E}_0)$ vanishes on the edges of Q. So, $\operatorname{div}(\mathbf{E}_0) \in H_0^{1/2}(Q)$. Moreover,

$$\begin{aligned} M(\mathbf{E}_0, \mathbf{H}_0) &= -(\mathbf{E}_0, \mathbf{H}_0) - \begin{pmatrix} 0 \\ 0 \\ 0 \\ \pi\sin(\pi x_1)\big(\cos(\pi x_2)\sin(\pi x_3) - \sin(\pi x_2)\cos(\pi x_3)\big) \\ \pi\sin(\pi x_2)\big(\sin(\pi x_1)\cos(\pi x_3) - \cos(\pi x_1)\sin(\pi x_3)\big) \\ \pi\sin(\pi x_3)\big(\cos(\pi x_1)\sin(\pi x_2) - \sin(\pi x_1)\cos(\pi x_2)\big) \end{pmatrix} \\ &=: -(\mathbf{E}_0, \mathbf{H}_0) - (\widetilde{\mathbf{E}}_0, \widetilde{\mathbf{H}}_0). \end{aligned}$$

We see $(\widetilde{\mathbf{E}}_0, \widetilde{\mathbf{H}}_0) \in D(M)$, which gives $(\mathbf{E}_0, \mathbf{H}_0) \in D(M^2)$.

Setting

$$\begin{aligned} (\widehat{\mathbf{E}}_0, \widehat{\mathbf{H}}_0) :&= M(\widetilde{\mathbf{E}}_0, \widetilde{\mathbf{H}}_0) \\ &= \pi^2 \begin{pmatrix} \sin(\pi x_3)\big(\cos(\pi x_1)\cos(\pi x_2) + \sin(\pi x_1)\sin(\pi x_2)\big) \\ \sin(\pi x_1)\big(\cos(\pi x_2)\cos(\pi x_3) + \sin(\pi x_2)\sin(\pi x_3)\big) \\ \sin(\pi x_2)\big(\cos(\pi x_1)\cos(\pi x_3) + \sin(\pi x_1)\sin(\pi x_3)\big) \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ &\quad - \pi^2 \begin{pmatrix} -\sin(\pi x_2)\big(\sin(\pi x_1)\sin(\pi x_3) + \cos(\pi x_1)\cos(\pi x_3)\big) \\ -\sin(\pi x_3)\big(\sin(\pi x_1)\sin(\pi x_3) + \cos(\pi x_1)\cos(\pi x_2)\big) \\ -\sin(\pi x_1)\big(\sin(\pi x_2)\sin(\pi x_3) + \cos(\pi x_2)\cos(\pi x_3)\big) \\ 0 \\ 0 \\ 0 \end{pmatrix}, \end{aligned}$$

we see

$$\mathrm{tr}_t(\widehat{\mathbf{E}}_0) = \begin{pmatrix} 0 \\ \mp\pi^2 \sin(\pi x_3)\cos(\pi x_2) \\ \mp\pi^2 \sin(\pi x_2)\cos(\pi x_3) \end{pmatrix} \neq 0$$

on $\Gamma_1^{\pm}$ (and analogously on $\Gamma_2^{\pm}$ and $\Gamma_3^{\pm}$). Due to $\mathrm{tr}_t(\mathbf{E}_0) = 0$ and $\mathrm{tr}_t(\widetilde{\mathbf{E}}_0) = 0$ on Γ we infer $\mathrm{tr}_t(M^2(\mathbf{E}_0, \mathbf{H}_0)) \neq 0$ on Γ. So, $(\mathbf{E}_0, \mathbf{H}_0) \notin D(M^3)$.

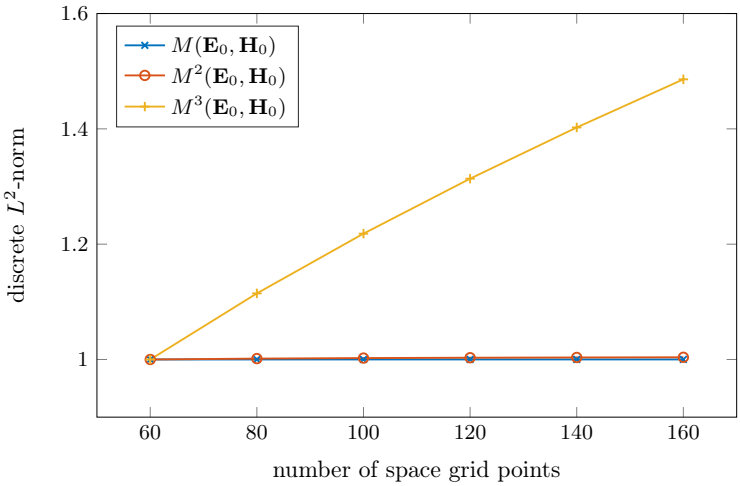


Figure 10.6.: Behaviour of the discrete L^2-norm of the initial function on different space grids when applying M.

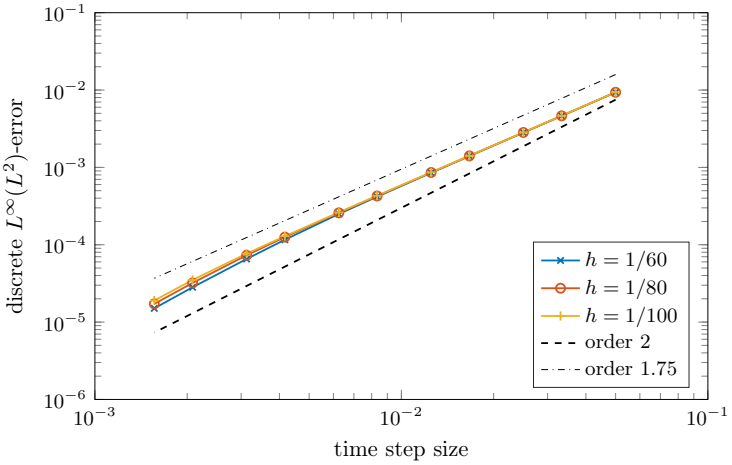


Figure 10.7.: Errors of the ADI method, using an initial function in $X_{\mathrm{div}}^{(2)} \setminus D(M_{\mathrm{div}}^{(2)})$ and the space mesh width h.

To illustrate these analytical investigations numerically we display in Figure 10.6 the discrete L^2-norm of $(\mathbf{E}_0, \mathbf{H}_0)$ on different space grids after applying M, M^2 and M^3. We

normalised the values by setting it to one on the coarsest space grid since only the relative increase is important.

The results of the ADI splitting, depicted in Figure 10.7, show a reduction of the convergence order to 1.75. Nevertheless, we see for small time step sizes an increase of the convergence order to two since we are then in the non-stiff regime.

Bibliography

[1] R. A. Adams and J. J. F. Fournier, *Sobolev Spaces*, Second edition, Elsevier, Oxford, 2003.

[2] H. Amann, *Linear and Quasilinear Parabolic Problems. Volume I: Abstract Linear Theory*, Basel, Birkhäuser, 1995.

[3] C. Amrouche, C. Bernardi, M. Dauge and V. Girault, *Vector potentials in three-dimensional non-smooth domains*, Math. Mathods Appl. Sci. **21**(9) (1998), 823–864.

[4] W. Auzinger, T. Kassebacher, O. Koch and M. Thalhammer, *Adaptive splitting methods for nonlinear Schrödinger equations in the semiclassical regime*, Numer. Algor. **72** (2016), 1–35.

[5] W. Auzinger, O. Koch and M. Thalhammer, *Defect-based local error estimators for splitting methods, with application to Schrödinger equations, Part I: The linear case*, J. Comput. Appl. Math. **236** (2012), 2643–2659.

[6] W. Auzinger, O. Koch and M. Thalhammer, *Defect-based local error estimators for splitting methods, with application to Schrödinger equations, Part II: Higher-order methods for linear problems*, J. Comput. Appl. Math. **255** (2014), 384–403.

[7] W. Auzinger, H. Hofstätter, O. Koch and M. Thalhammer, *Defect-based local error estimators for splitting methods, with application to Schrödinger equations, Part III: The nonlinear case*, J. Comput. Appl. Math. **273** (2015), 182–204.

[8] A. Benyi and T. Oh, *The Sobolev inequality on the torus revisited*, Publ. Math. Debrecen **83** (2013), 359–374.

[9] J. Bergh and J. Löfström, *Interpolation Spaces*, Springer, 1976.

[10] C. Besse, B. Bidégaray and S. Descombes, *Order estimates in time of splitting methods for the nonlinear Schrödinger equation*, SIAM J. Numer. Anal. **40** (2002) 26–40.

[11] J. Bourgain, *Global Solutions of Nonlinear Schrödinger Equations*, American Mathematical Society, Colloquium Publications **46**, 1999.

[12] M. Caliari, G. Kirchner and M. Thalhammer, *Convergence and energy conservation of the Strang time-splitting hermite spectral method for nonlinear Schrödinger equations*. Unpublished manuscript (2007), `http://profs.sci.univr.it/ caliari/pdf/preCKT07.pdf`.

[13] T. Cazenave, *Semilinear Schrödinger Equations*, American Mathematical Society, Providence (RI), 2003.

[14] W. Chen, X. Li and D. Liang, *Energy-conserved splitting FDTD methods for Maxwell's equations*, Numer. Math. **108** (2008), 445–485.

[15] M. Costabel and M. Dauge, *Singularities of electromagnetic fields in polyhedral domains*, Arch. Rational Mech. Anal. **151** (2000), 221–276.

[16] R. Dautray and J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology, Volume 3: Spectral Theory and Applications*, Springer, Berlin, reprint, 2000.

[17] E. B. Davies, *Spectral Theory and Differential Operators*, Cambridge Univ. Press, 1995.

[18] J. Douglas, *On the numerical integration of* $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial u}{\partial t}$ *by implicit methods*, J. Soc. Indust. Appl. Math. **3**(1) (1955), 42–66.

[19] J. Douglas and H. H. Rachford, *On the numerical solution of heat conduction problems in two and three space variables*, Trans. Numer. Math. Soc. **82** (1956), 421-439.

[20] L. Einkemmer and A. Ostermann, *Overcoming order redcution in diffusion-reaction splitting. Part 1: Dirichlet boundary conditions*, SIAM J. Sci. Comp. **37**(3) (2015), 1577–1592.

[21] L. Einkemmer and A. Ostermann, *Overcoming order redcution in diffusion-reaction splitting. Part 2: Oblique boundary conditions*, SIAM J. Sci. Comp. **38**(6) (2016), 3741–3757.

[22] J. Eilinghoff, R. Schnaubelt and K. Schratz, *Fractional error estimates of splitting schemes for the nonlinear Schrödinger equation*, JMAA **442** (2016), 740–760.

[23] K.-J. Engel and R. Nagel, *One-parameter Semigroups for Linear Evolution Equations*, In: Graduate texts in mathematics, Springer, New York, 2000.

[24] L. C. Evans, *Partial Differential Equations*, In: Graduate Studies in Mathematics, Vol. 19, AMS, Providence, Rhode Island, 2nd ed., 2010.

[25] E. Faou, *Geometric Numerical Integration of Schrödinger Equations*, European Mathematical Society, Zürich, 2002.

[26] L. Gauckler, *Numerical long-time energy conservation for the nonlinear Schrödinger equation*, IMA Journal of Numerical Analysis **00** (2014), 1–24.

[27] L. Gauckler, *Error analysis of trigonometric integrators for semilinear wave equations*, SIAM J. Numer. Anal. **53** (2015), 1082–1106.

[28] L. Gauckler and C. Lubich, *Nonlinear Schrödinger equations and their spectral semi-discretizations over long times*. Found. Comput. Math. **10** (2010), 141–169.

[29] L. Gauckler and C. Lubich, *Splitting integrators for nonlinear Schrödinger equations over long times*. Found. Comput. Math. **10** (2010), 275–302.

[30] S. D. Gedney, G. Liu, J. A. Roden and A. Zhu, *Perfectly matched layer media with CFS for an unconditionally stable ADI-FDTD method*, IEEE Transactions on antennas and propagation **49**(11) (2001), 1554–1559.

[31] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer, 2001 (reprint of the 1998 edition).

[32] P. Grisvard, *Spazi di tracce e applicazioni*, Rend. Math. **5** (1973), 657–729.

[33] E. Hairer, C. Lubich and G. Wanner, *Geometric Numerical Integration: Structure-preserving Algorithms for Ordinary Differential Equations*, 2nd edition, Springer, 2006.

[34] E. Hairer, S. P. Nørset and G. Wanner, *Solving Ordinary Differential Equations I. Nonstiff Problems*, 2nd edition, Springer, 1993.

[35] E. Hansen and A. Ostermann, *Dimension splitting for evolution equations*, Numer. Math. **108** (2008), 557-570.

[36] E. Hansen and A. Ostermann, *Exponential splitting for unbounded operators*, Math. Comp. **78**(267) (2009), 1485–1496.

[37] M. Hochbruck, T. Jahnke and R. Schnaubelt, *Convergence of an ADI splitting for Maxwell's equations*, Numer. Math. **129**(3) (2014), 535-561.

[38] M. Hochbruck and A. Ostermann, *Exponential integrators*, Acta Numerica (2010), 209–286.

[39] M. Hochbruck and A. Sturm, *Error analysis of a second-order locally implicit method for linear Maxwell's equations*, SIAM J. Numer. Anal. **54**(5) (2016), 3167-3191.

[40] H. Holden, C. Lubich and N. H. Risebro, *Operator splitting for partial differential equations with Burgers' nonlinearity*, Math. Comp. **82** (2013), 173–185.

[41] R. Holland, *Implicit three-dimensional finite differencing of Maxwell's equations*, IEEE Transactions on Nuclear Science **31**(6) (1984), 1322–1326.

[42] L. I. Ignat, *A splitting method for the nonlinear Schrödinger equation*, J. Differential Equations **250** (2011), 3022–3046.

[43] T. Jahnke and C. Lubich, *Error bounds for exponential operator splitting*, BIT **40** (2000), 735–744.

[44] T. Jahnke, M. Mikl and R. Schnaubelt, *Strang splitting for a semilinear Schrödinger equation with damping and forcing*, J. Math. Anal. Appl. (2015), accepted.

[45] D. Jerison and C. E. Kenig, *The inhomogeneous Dirichlet problem in Lipschitz domains*, Journal of Functional Analysis **130** (1995), 161–219.

[46] T. Kato, *On nonlinear Schrödinger equations, II. H^s-solutions and unconditional well-posedness*, J. Anal. Math. **67** (1995), 281–306.

[47] T. Kato, *Perturbation Theory for Linear Operators*, Springer Verlag Heidelberg, 2nd edition, reprint of the 1980 edition, 1995.

[48] P. C. Kunstmann and L. Weis, *Maximal L^p-regularity for Parabolic Equations, Fourier Multiplier Theorems and H^∞-functional Calculus* in Functional Analytic Methods for Evolution Equations, Springer, Berlin, 2004.

[49] G. Liu and S. D. Gedney, *Perfectly matched layer media for an unconditionally stable three-dimensional ADI-FDTD method*, IEEE Microwave and guided wave letters **10**(7) (2000), 261–263.

[50] J. Lu and J. L. Marzuola, *Strang splitting methods for a quasilinear Schrödinger equation: convergence, instability, and dynamics*, Commun. Math. Sci. **13** (2015), 1051–1074.

[51] C. Lubich, *On splitting methods for Schrödinger-Poisson and cubic nonlinear Schrödinger equations*, Math. Comp. **77** (2008), 2141–2153.

[52] A. Lunardi, *Interpolation Theory*, Edizione della Normale, Pisa, 2009.

[53] G. I. Marchuk, *Some application of splitting-up methods to the solution of mathematical physics problems*, Aplikace matematiky, **13**(2) (1968), 103–132.

[54] R. I. McLachlan and G. R. W. Quispel, *Splitting methods*, Acta Numerica (2002), 341–434.

[55] J. Moloney and A. Newell, *Nonlinear Optics*, Westview Press, Boulder (Co), 2004.

[56] D. Müller, *Well-posedness for a General Class of Quasilinear Evolution Equations – With Applications to Maxwell's Equations*, PhD thesis, 2014.

[57] T. Namiki, *A new FDTD algorithm based on alternating-direction implicit method*, IEEE Transactions on microwave theory and techniques **47**(10) (1999), 2003–2007.

[58] J. Nečas, *Direct Methods in the Theory of Elliptic Equations*, Springer, Heidelberg, 2012.

[59] A. Ostermann and K. Schratz, *Error analysis of splitting methods for inhomogeneous evolution equations*, Appl. Num. Math. **62** (2012), 1436–1446.

[60] A. Ostermann and K. Schratz, *Low regularity exponential-type integrators for semilinear Schrödinger equations in the energy space*, https://arxiv.org/abs/1603.07746, (2016).

[61] E. Ouhabaz, *Analysis of Heat Equations in Domains*, London Mathematical Society monographs series **31**, Princeton, NJ, 2005.

[62] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer, New York, 1992.

[63] D. W. Peaceman and H. H. Rachford, *The numerical solution of parabolic and elliptic differential equations*, J. Soc. Indust. Appl. Math. **3**(1) (1955), 28–41.

[64] M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis*, Academic Press, San Diego, 2nd edition, 1980.

[65] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, 3rd Ed. Springer, New York, 2002.

[66] G. Strang, *On the construction and comparison of different schemes*, SIAM J. Numer. Anal. (1968), 506–512.

[67] C. Sulem and P.-L. Sulem, *Nonlinear Schrödinger Equation: Self-focusing and Wave Collapse*, Springer, 1999.

[68] M. Thalhammer, *Higher-order exponential operator splitting methods for time-dependent Schrödinger equations*, SIAM J. Numer. Anal. **46** (2008), 2022–2038.

[69] M. Thalhammer, M. Caliari and C. Neuhauser, *High-order time-splitting Hermite and Fourier spectral methods*, J. Comput. Phys. **228** (2009), 822–832.

[70] H. Triebel, *Interpolation Theory, Function Spaces, Differential Operators*, Deutscher Verlag der Wissenschaften, Berlin, 1978.

[71] H. F. Trotter, *On the product of semigroups of operators*, Proceedings of the American Mathematical Society, **10**(4) (1959), 545–551.

[72] K. S. Yee, *Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media*, IEEE Transaction on antennas and propagation **14**(3) (1966), 302–307.

[73] H. Yoshida, *Construction of higher order symplectic integrators*, Physics Letters, **150** (1990), 262–268.

[74] F. Zheng and Z. Chen, *Numerical dispersion analysis of the unconditionally stable 3-D ADI-FDTD method*, IEEE Transactions on microwave theory and techniques **49**(5) (2001), 1006–1009.

[75] F. Zheng, Z. Chen and J. Zhang, *Toward the development of a three-dimensional unconditionally stable finite-difference time-domain method*, IEEE Transaction on microwave theory and techniques **48**(9) (2000), 1550–1558.

[76] F. Zheng, Z. Chen and J. Zhang, *A finite-difference time-domain method without the Courant stability condition*, IEEE Microwave and guided wave letters **9**(11) (1999), 441–443.

Index

List of Symbols

This list of symbols is ordered by their appearance in the text and grouped by the parts of this thesis.

Part I

symbol	meaning
I	the identity operator
$\mathbb{1}$	the function being constant one
$\mathbb{1}_A$	the indicator function of a set A
$\hookrightarrow$	continuous embedding
$X \cong Y$	Y is isomorphic to X with equivalent norm
$\langle \cdot, \cdot \rangle$	a duality paring
$\mathcal{B}(X, Y)$	the set of linear and bounded operator from X to Y
$\mathcal{B}(X)$	the set of linear and bounded operator from X to X
$\|\cdot\|_{D(A)}$	the graph norm with respect to the operator A
$(\lambda I - A)^{-1}$	the resolvent of A for λ in the resolvent set
Ω	an open or Borel measurable subset of $\mathbb{R}^d$ (with $d \in \mathbb{N}$)
C_c^∞	the set of infinitely often differentiable functions with compact support
∂_j	the partial or weak derivative with respect to the j-th variable
L^1_{loc}	the space of locally integrable functions
$L^p, \quad p \in [1, \infty]$	the Lebesgue spaces
$W^{k,p}, \quad k \in \mathbb{N}_0$	the Sobolev spaces
$W^{s,p}, \quad s \geq 0$	the (fractional) Sobolev spaces
$(X, Y)_{\eta,2}$, $\eta \in (0, 1)$	real interpolation space with the parameters η and 2
$H^s, \quad s \geq 0$	the (fractional) Sobolev spaces (with respect to L^2)
$\mathcal{F}$ and $\mathcal{F}^{-1}$	the Fourier transform and its inverse
$\|\cdot\|_{X \cap Y}$	$\|\cdot\|_X + \|\cdot\|_Y$
$\mathbb{T}^d$	the d-dimensional torus
$\simeq$	equal up to a multiplicative constant

X_{-1}^A	the Sobolev space of order -1 associated to the semigroup generated by A. see (1.8)
$(X, \|\cdot\|_Y)^\sim$	completion of X with respect to the norm $\|\cdot\|_Y$

Part II

symbol	meaning
μ	the sign of the nonlinearity of the NLS
Ω	domain of interest, either $\mathbb{R}^d$ or $\mathbb{T}^d$
∂_t	the partial derivative with respect to time
Δ	the Laplace operator
Φ_τ	the Lie splitting scheme for the cubic NLS with time step size τ, see (2.2)
$T(\cdot)$	the free Schrödinger group
Ψ_τ	the Strang splitting scheme for the cubic NLS with time step size τ, see (2.3)
A and B	the splitting operators for the cubic NLS, see (2.4)
M_s	the supremum norm of the solution of the cubic NLS in H^s over $[0, T]$
$C^{0,\theta}$	the space of θ-Hölder continuous functions
$C^{1,\theta}$	the space of differentiable functions whose derivative is θ-Hölder continuous
$D^2 f$	the matrix of the second-order (weak) derivatives of a function f

Part III

symbol	meaning
Ω	a cuboid in $\mathbb{R}^3$
ε	the electric permittivity
σ	the electric conductivity
$\mathbf{J}_0$	the electric current density
μ	the magnetic permeability
ρ	the electric charge density
C_1 and C_2	the parts of the split curl-operator, see (7.22)
A and B	the splitting operators for the ADI scheme, see (7.23)
Γ	the boundary of the cuboid Ω

$\Gamma_j, j = 1, 2, 3$	the faces of Γ that are orthogonal to the respectively coordinate axis
X	$L^2(\Omega)^6$
tr	the Dirichlet trace
tr_t and tr_n	the tangential and the normal trace
$H_0(\Omega, \mathrm{curl})$	the space of functions in $H(\Omega, \mathrm{curl})$ with zero tangential trace
M	the Maxwell operator, see (7.3)
X_0, $X_{\mathrm{div}}^{(0)}$ and $X_{\mathrm{div}}^{(2)}$	subspaces of X, see (7.4) and (7.6)
$(M(u, v))_{1/2}$	the components one till three / four till six of $M(u, v)$
M_0, $M_{\mathrm{div}}^{(0)}$ and $M_{\mathrm{div}}^{2)}$	restrictions of the Maxwell operator to X_0, $X_{\mathrm{div}}^{(0)}$ and a subspace of $X_{\mathrm{div}}^{(2)}$, see (7.7)
$[X, Y]_\eta$	complex interpolation space with parameter η
A_0 and B_0	the splitting operators with zero conductivity
$H^{\theta,p}$, $H_0^{\theta,p}$	Bessel potential spaces
Y	a subspace of $H^1(\Omega)^6$ with certain boundary conditions, see (7.29)
A_Y and B_Y	the part of A and B in Y
Z	a subspace of $H^2(\Omega)^6$ with certain boundary conditions, see (7.30)
A_Z and B_Z	a restriction of A and B to a subspace of Z, see (7.32) and (7.33)