Error analysis for space and time discretizations of quasilinear wave-type equations

Zur Erlangung des akademischen Grades

DOKTOR DER NATURWISSENSCHAFTEN

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

DISSERTATION

von

Bernhard Maier

Tag der mündlichen Prüfung: 27. Mai 2020

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

Acknowledgement

At this point I would like to express my gratitude to all those who contributed directly or indirectly to this thesis.

First of all, I thank my supervisor Prof. Dr. Marlis Hochbruck for the opportunity to work on this exciting research project. It was her confidence in my abilities and her extensive support that made this thesis possible. Furthermore, I would like to thank my second supervisor Prof. Dr. Roland Schnaubelt for always having an open door for me. I am also very grateful to Prof. Dr. Charalambos Makridakis for his report as well as his expertise.

This thesis was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 258734477 – SFB 1173. Besides, I highly appreciate the diverse range of possibilities offered by the CRC and the iRTG, both professionally and personally.

Moreover, I would like to thank the extended numerical analysis group for the pleasant working environment including the welcome distractions. In particular, I would like to mention Benjamin, Constantin and Jan, who have been incredibly patient in proofreading this work.

I am very grateful to my family for their unconditional and kind support. None of this would have been possible without their encouragement.

My deepest gratitude goes to Joana, the love of my life.

Contents

CHAPTER 1

Introduction

Motivation

Wave-type problems are fundamental models in physics, in particular in acoustics and optics. For example, the acoustic wave equation is the basic model to describe the propagation of sound waves. Furthermore, the propagation of light is governed by the Maxwell equations.

For a long time, scientists working in these fields focused on linearized models to describe the fundamental relations between these waves and the surrounding media. However, these linear models have their limits, especially with respect to the frequency and intensity of the waves considered. To overcome these limits, the research fields of nonlinear acoustics and nonlinear optics emerged in the early 20th century. Since then, nonlinear wave-type equations became more and more important, as various nonlinear models were introduced. With respect to our previous examples, this includes for instance the Westervelt equation to describe the propagation of ultrasound as well as nonlinear constitutive relations for the Maxwell equations to model the Kerr effect.

From the mathematical perspective, these nonlinear problems differ essentially from their linear counterparts. However, there exists a particular class of nonlinear problems with an underlying linear structure, namely quasilinear problems. Moreover, since the analytical theory for linear wave-type problems is well established, quasilinear wave-type problems proved to be a good starting point to investigate the wellposedness of nonlinear problems. In this context, a major achievement was the extension of semigroup theory to quasilinear wave-type problems in [Kato, 1975], which yields wellposedness for a very general class of quasilinear wave-type problems on the full space including various important applications. In subsequent papers these ideas were extended to specific quasilinear wave-type problems on regular bounded domains. In addition, since Kato's framework is restrictive concerning the boundary conditions, alternative approaches using for example energy techniques were developed to prove wellposedness for specific wave-type problems on bounded domains.

Regarding the initial examples, the wellposedness of the undamped Westervelt equation subject to homogeneous Dirichlet boundary conditions was proven in [Dörfler et al., 2016] based on Kato's framework. Furthermore, the wellposedness of the Maxwell equations with Kerr nonlinearity and perfectly conducting boundary conditions is shown in [Spitz, 2019] using energy techniques.

In contrast to the analytical wellposedness theory, the numerical analysis of quasilinear wave-type problems is much less developed. Nevertheless, there are a few results for the discretization of quasilinear second-order wave-type problems. For example, the full discretization with linearized implicit one- and two-step time-integration schemes is considered in [Ewing, 1980], [Bales, 1986], [Bales, 1988], [Bales and Dougalis, 1989], and [Makridakis, 1993]. This includes the construction of the schemes as well as rigorous error estimates. Moreover, based on Banach's fixed-point theorem, [Makridakis, 1993] provides error estimates for the space discretization with finite elements and the full discretization with a class of fully implicit two-step time-integration schemes.

Besides these rather general results, there are also more specific results: For example, the space discretization of quasilinear, elastic wave equations with a discontinuous Galerkin discretization is investigated in [Ortner and Süli, 2007]. Furthermore, based on the Petrov-Galerkin method, the full discretization of a specific class of quasilinear second-order wave equations in 1D is considered in [Gerner, 2013]. Based on stability assumptions on the numerical solution, the author proves first order convergence for the full discretization with a variant of the implicit midpoint rule. More recently, in [Gauckler et al., 2019] the authors consider the full discretization of a special class of quasilinear second-order wave equations in 1D with constant coefficients and periodic boundary conditions using explicit trigonometric integrators. However, note that these schemes require the evaluation of matrix functions.

With respect to the initial examples, on the one hand, there are two recent results for the Westervelt equation with strong damping, i.e., the space discretization with finite elements and discontinuous Galerkin finite elements is analyzed in [Nikolić and Wohlmuth, 2019] and [Antonietti et al., 2020], respectively. On the other hand, up to our knowledge the error analysis for a discretization of the quasilinear Maxwells equations was not considered so far. Nevertheless, in [Pototschnig et al., 2009] an exponential integrator for the quasilinear Maxwells equations was proposed and tested numerically.

Just recently, the time discretization of a very general class of quasilinear wave-type problems with algebraically stable Runge–Kutta schemes was considered in [Hochbruck and Pažur, 2017], [Hochbruck et al., 2018], and [Kovács and Lubich, 2018]. Based on Kato's framework, the authors prove both wellposedness and error estimates. In addition, a unified error analysis for nonconforming space and time discretizations of linear and semilinear wave-type problems was presented in [Hipp et al., 2019] and [Hochbruck and Leibold, 2019], respectively. These two concepts form the basis of this thesis.

Main results

In this thesis we present a rigorous error analysis for the abstract space and time discretization of a very general class of quasilinear wave-type problems. Up to our knowledge this includes the first error analysis for first-order quasilinear wave-type problems.

Concerning the error analysis for the space discretization, we employ semigroup theory to prove wellposedness as well as an error estimate for the spatially discrete problem. Compared to previous results which are mostly based on Banach's fixed-point theorem, this approach provides better insight into the individual error contributions. Furthermore, since wellposedness results for quasilinear wave-type problems are in general based on severe regularity assumptions with respect to the boundary of the domain, we consider nonconforming space discretizations in order to allow for domain approximation.

Based on these results for the space discretization, we further prove wellposedness as well as rigorous error estimates for the full discretization with three different one-step time-integration schemes. On the one hand, we consider the implicit midpoint rule and a linearized version thereof. On the other hand, we also investigate the leapfrog scheme, which is an explicit scheme. We emphasize that, up to our knowledge, the full discretization with both explicit and nonlinear implicit time-integration schemes was prior to this thesis only analyzed for special classes of quasilinear second-order wave-type problems in 1D, as mentioned above.

Throughout this thesis, we illustrate the relevance of the abstract framework by application of our results to the undamped Westervelt equation and the Maxwell equations with Kerr nonlinearity. Finally, we conclude with numerical examples.

Outline

This thesis is organized as follows. In Chapter 2 we present the basic notation and some mathematical tools which are used throughout this thesis.

In Chapter 3 we introduce the abstract quasilinear wave-type problem (3.1) and state basic assumptions. Furthermore, we introduce the Westervelt equation and the Maxwell equations with Kerr nonlinearity as specific examples, which are revisited frequently throughout this thesis.

Chapter 4 is denoted to the analysis of nonconforming space discretizations of quasilinear wave-type problems. To this end, we first establish the general framework for the nonconforming space discretization, including assumptions on the discrete operators. After a brief overview to semigroup theory for nonautonomous Cauchy problems, we derive an abstract error estimate. Based on these results, we provide error estimates for the space discretization of the specific examples in Chapter 5.

In Chapter 6 we briefly introduce the leapfrog scheme as well as algebraically stable Runge–Kutta schemes, with a special focus on the implicit midpoint rule. We further review recent results for the time discretization of abstract quasilinear wave-type problems.

The full discretization of quasilinear wave-type problems with two variants of the implicit midpoint rule and the leapfrog scheme is considered in Chapter 7. First, we prove wellposedness and a rigorous error estimate for a linearized version of the implicit midpoint rule. We then extend these results to the implicit midpoint rule using fixed-point iterations. Furthermore, based on an alternative error analysis for the linearized variant, we prove wellposedness and a rigorous error estimate for the leapfrog scheme. In Chapter 8 these results are applied to the specific examples.

For the Westervelt equation in one and two space dimensions, we validate these theoretical results in Chapter 9 with numerical experiments.

Finally, we close this thesis with a brief summary of the main results and concluding remarks in Chapter 10.

CHAPTER 2

Preliminaries

Throughout this thesis, we use the following notation.

Miscellaneous We consider a domain $\Omega \subset \mathbb{R}^d$ with spatial dimension $d \in \mathbb{N}$. For $T > 0$ we denote the time interval by $J_T := [0, T]$. Furthermore, for a normed vector space $\mathcal{H} = (\mathcal{H}, \|\cdot\|_{\mathcal{H}})$ and $R > 0$, we define

$$B_{\mathcal{H}}(R) := \{\varphi \in \mathcal{H} \mid \|\varphi\|_{\mathcal{H}} < R\},$$

i.e., the open sphere in $\mathcal{H}$ centered at 0 with radius R. Finally, we use a generic constant $C > 0$, which may have different values at different occurrences.

Normed vector spaces For normed vector spaces $\mathcal{H} = (\mathcal{H}, \|\cdot\|_{\mathcal{H}})$ and $\mathcal{G} = (\mathcal{G}, \|\cdot\|_{\mathcal{G}})$, we denote the set of all bounded linear operators mapping from $\mathcal{H}$ to $\mathcal{G}$ by $\mathcal{L}(\mathcal{H}, \mathcal{G})$ and simply write $\mathcal{L}(\mathcal{H})$ if $\mathcal{H} = \mathcal{G}$. The operator norm is given by

$$\|A\|_{\mathcal{L}(\mathcal{H},\mathcal{G})} := \sup_{x \in \mathcal{H}\setminus\{0\}} \frac{\|Ax\|_{\mathcal{G}}}{\|x\|_{\mathcal{H}}}, \qquad A \in \mathcal{L}(\mathcal{H}, \mathcal{G}).$$

Furthermore, we denote the identity operator by $\mathrm{Id} \in \mathcal{L}(\mathcal{H})$, where we do not explicitly specify the space $\mathcal{H}$ for the sake of presentation. Finally, the norm of a product space $\mathcal{X} = \mathcal{H} \times \mathcal{G}$ is given by

$$\|\xi\|_{\mathcal{X}}^2 := \|\varphi\|_{\mathcal{H}}^2 + \|\psi\|_{\mathcal{G}}^2, \qquad \xi = \begin{pmatrix} \varphi \\ \psi \end{pmatrix} \in \mathcal{X}.$$

Vector algebra For $a, b, c \in \mathbb{R}^d$ with $a = (a_1, \dots, a_d)$ and $b = (b_1, \dots, b_d)$ we write

$$a \cdot b := \sum_{i=1}^{d} a_i b_i, \qquad a \otimes b := \begin{pmatrix} a_1 b_1 & \cdots & a_1 b_d \\ \vdots & \ddots & \vdots \\ a_d b_1 & \cdots & a_d b_d \end{pmatrix}$$

for the inner and outer product of a and b, respectively. These vector products satisfy

$$(a \otimes b)c = a(b \cdot c) \tag{2.1}$$

as well as the bounds

$$\|(a \cdot b)c\|_2 \leq d\|a\|_\infty \|b\|_\infty \|c\|_2, \qquad \|(a \otimes b)c\|_2 \leq d\|a\|_\infty \|b\|_\infty \|c\|_2, \tag{2.2}$$

for $\|\cdot\|_2$ and $\|\cdot\|_\infty$ denoting the Euclidean norm and the maximum norm, respectively. Furthermore, for $d = 3$ the cross product of a and b, i.e.,

$$a \times b \coloneqq \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix},$$

satisfies

$$\|a \times b\|_2 \leq \|a\|_2 \|b\|_2. \tag{2.3}$$

Differential operators Let $\varphi : J_T \times \Omega \to \mathbb{R}$ and $\psi : \Omega \to \mathbb{R}^3$ be sufficiently smooth functions. We denote the partial derivative of φ with respect to time by $\partial_t \varphi$. Concerning the spatial derivatives, we denote the gradient of φ by $\nabla \varphi$ with $\nabla = (\partial_1, \dots, \partial_d)$. Moreover, the divergence and the curl of ψ are given by $\nabla \cdot \psi$ and $\nabla \times \psi$, respectively. Finally, we denote the laplacian of φ by $\Delta \varphi = \nabla \cdot \nabla \varphi$. Note that in the special case $d = 1$, we simply write $\partial_x \varphi$ instead of $\nabla \varphi$.

Function spaces We denote the standard Lebesgue space of real-valued functions by $L^2(\Omega)$, equipped with the inner product

$$(\varphi \mid \psi)_{L^2(\Omega)} \coloneqq \int_\Omega \varphi(x)\psi(x)\,\mathrm{d}x, \qquad \varphi, \psi \in L^2(\Omega),$$

and the corresponding norm $\|\cdot\|_{L^2(\Omega)}$. Moreover, $L^\infty(\Omega)$ denotes the space of all essentially bounded measurable functions with

$$\|\varphi\|_{L^\infty(\Omega)} \coloneqq \operatorname*{ess\,sup}_{x \in \Omega} |\varphi(x)|, \qquad \varphi \in L^\infty(\Omega).$$

For vector-valued versions of these function spaces, we define for $p \in \{2, \infty\}$ the norms

$$\|\varphi\|_{L^p(\Omega)^d} \coloneqq \big\| \|\varphi\|_p \big\|_{L^p(\Omega)}, \qquad \varphi \in L^p(\Omega)^d. \tag{2.4}$$

For a multiindex $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$ and $k \in \mathbb{N}$ we set $\partial_\alpha \varphi \coloneqq \partial_1^{\alpha_1} \dots \partial_d^{\alpha_d} \varphi$ for $\varphi : \Omega \to \mathbb{R}$ being sufficiently smooth. Then, the Sobolev space of order k is defined as

$$H^k(\Omega) \coloneqq \{\varphi \in L^2(\Omega) \mid \partial_\alpha \varphi \in L^2(\Omega), |\alpha| \leq k\},$$

with $|\alpha| \coloneqq \sum_{i=1}^d \alpha_i$. Note that equipped with the inner product

$$(\varphi \mid \psi)_{H^k(\Omega)} \coloneqq \sum_{|\alpha| \leq k} (\partial_\alpha \varphi \mid \partial_\alpha \psi)_{L^2(\Omega)}, \qquad \varphi, \psi \in H^k(\Omega),$$

$H^k(\Omega)$ is a Hilbert space. Furthermore, we define the corresponding norm

$$\|\varphi\|_{H^k(\Omega)}^2 \coloneqq \sum_{|\alpha|\le k} \|\partial_\alpha \varphi\|_{L^2(\Omega)}^2, \qquad \varphi \in H^k(\Omega),$$

as well as the seminorm

$$|\varphi|_{H^k(\Omega)}^2 \coloneqq \sum_{|\alpha|= k} \|\partial_\alpha \varphi\|_{L^2(\Omega)}^2, \qquad \varphi \in H^k(\Omega).$$

Moreover, for $C_0^\infty(\Omega)$ being the space of all compactly supported smooth functions on Ω, we define the space $H_0^1(\Omega)$ to be the closure of $C_0^\infty(\Omega)$ with respect to $|\cdot|_{H^1(\Omega)}$. We define

$$(\varphi \mid \psi)_{H_0^1(\Omega)} \coloneqq (\nabla \varphi \mid \nabla \psi)_{L^2(\Omega)^d}\,, \qquad \varphi, \psi \in H_0^1(\Omega),$$

which yields $\|\varphi\|_{H_0^1(\Omega)} = |\varphi|_{H^1(\Omega)}$.

Furthermore, we set

$$H(\mathrm{curl}, \Omega) \coloneqq \{\varphi \in L^2(\Omega)^3 \mid \nabla\times \varphi \in L^2(\Omega)^3\},$$

with norm

$$\|\varphi\|_{H(\mathrm{curl},\Omega)}^2 \coloneqq \|\varphi\|_{L^2(\Omega)^3}^2 + \|\nabla\times\varphi\|_{L^2(\Omega)^3}^2,$$

Again, $H_0(\mathrm{curl}, \Omega)$ denotes the closure of $C_0^\infty(\Omega)^3$ with respect to $\|\cdot\|_{H(\mathrm{curl},\Omega)}$.

Sobolev embedding Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with C^1-boundary. Furthermore, let $\ell \in \mathbb{N}$ with $\ell > \frac{d}{2}$. Then, we have

$$H^\ell(\Omega) \hookrightarrow C(\overline{\Omega}) \tag{2.5a}$$

as well as the Sobolev inequality

$$\|\varphi\|_{L^\infty(\Omega)} \le C_{\mathrm{S}} \|\varphi\|_{H^\ell(\Omega)}, \qquad \varphi \in H^\ell(\Omega), \tag{2.5b}$$

with a constant $C_{\mathrm{S}} > 0$ depending only on ℓ, d, and Ω, cf. [Evans, 2010, Sec. 5.6.3].

Trace inequality Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with Lipschitz boundary. Then, we have the trace inequality

$$\|\varphi\|_{L^2(\partial\Omega)} \le C_{\mathrm{tr}} \|\varphi\|_{L^2(\Omega)}^{\frac{1}{2}} \|\varphi\|_{H^1(\Omega)}^{\frac{1}{2}}, \qquad \varphi \in H^1(\Omega), \tag{2.6}$$

with a constant $C_{\mathrm{tr}} > 0$, cf. [Brenner and Scott, 2008, Thm. 1.6.6].

CHAPTER 3

Quasilinear evolution equations and specific examples

We consider quasilinear Cauchy problems of the form

$$\begin{cases} \Lambda(y(t))\partial_t y(t) = \mathrm{A}y(t) + \mathrm{F}(t, y(t)), & t \in J_T, \\ \qquad\qquad y(0) = y_0, \end{cases} \tag{3.1}$$

where Λ is a sufficiently regular nonlinear operator, A is a linear differential operator, F is a sufficiently regular nonlinear right-hand side, and $J_T = [0, T]$, $T < \infty$, denotes the time interval.

Although we are not aware of a wellposedness result for this general class of quasilinear problems, there are several results for specific examples falling into this framework; e.g., [Kato, 1975] and [Hughes et al., 1976] consider problems posed on the full space, whereas the authors of [Chen and von Wahl, 1982], [Dafermos and Hrusa, 1985], [Kato, 1985], [Koch, 1993], [Müller, 2014], [Dörfler et al., 2016], and [Spitz, 2019] consider problems posed on bounded domains. Nevertheless, we consider the discretization of the general problem (3.1), as this covers all these examples. In Section 3.1 we collect the assumptions on the operators appearing in (3.1), which are used throughout this thesis. In the following sections we then focus on specific examples and show that they fit in the general framework. The wellposedness of the Westervelt equation is presented in Section 3.2 based on results from [Dörfler et al., 2016]. Furthermore, we review in Section 3.3 the wellposedness result for quasilinear Maxwell equations from [Spitz, 2019].

3.1 General setting

As already stated in the introductory part of this chapter, we are not aware of a wellposedness result for the general problem (3.1). Nevertheless, we state in this section basic properties of the involved operators in order to transform the problem into an equivalent form, which is suitable for the analysis in the following chapters. However, these basic properties are not sufficient to guarantee wellposedness, as can be seen in the following sections for specific examples fitting into the general framework.

We start by collecting some basic properties of the operators Λ, A, and F that are used throughout this work.

Assumption 3.1. *There are Hilbert spaces $(\mathcal{X}, (\cdot \mid \cdot)_{\mathcal{X}})$ and $(\mathcal{Y}, (\cdot \mid \cdot)_{\mathcal{Y}})$ such that $\mathcal{Y} \hookrightarrow \mathcal{X}$, with a dense and continuous embedding. Further, there exists $R > 0$ such that the following properties hold.*

(Λ) *$\{\Lambda(\xi) \mid \xi \in B_{\mathcal{Y}}(R)\} \subset \mathcal{L}(\mathcal{X})$ is a family of symmetric operators, which are uniformly positive definite and bounded, i.e., there are constants $c_\Lambda, C_\Lambda > 0$ such that*

$$c_\Lambda \|\varphi\|_{\mathcal{X}}^2 \leq (\Lambda(\xi)\varphi \mid \varphi)_{\mathcal{X}}, \qquad \|\Lambda(\xi)\|_{\mathcal{L}(\mathcal{X})} \leq C_\Lambda, \qquad \varphi \in \mathcal{X},\, \xi \in B_{\mathcal{Y}}(R) \tag{3.2}$$

holds.

(A) $\mathrm{A} \in \mathcal{L}(D(\mathrm{A}), \mathcal{X})$ *with* $\mathcal{Y} \subset D(\mathrm{A})$ *and* $\overline{D(\mathrm{A})} = \mathcal{X}$, *where* $D(\mathrm{A})$ *denotes the domain of* A.

(F) $\mathrm{F} : J_T \times B_{\mathcal{Y}}(R) \to \mathcal{X}$ *is continuous in time and bounded, i.e., there is a constant* $C_\mathrm{F} > 0$ *such that* F *satisfies*

$$\|\mathrm{F}(t,\xi)\|_{\mathcal{X}} \leq C_\mathrm{F}, \qquad t \in J_T,\, \xi \in B_{\mathcal{Y}}(R).$$

Moreover, we emphasize that we state the assumptions on Λ and F only for spheres $B_{\mathcal{Y}}(R)$ in order to keep the notation simple. However, all our results can be generalized to bounded domains.

In the following, we always use R as the radius from Assumption 3.1. If (3.2) holds, the family of inverse operators $\{\Lambda(\xi)^{-1} \mid \xi \in B_{\mathcal{Y}}(R)\} \subset \mathcal{L}(\mathcal{X})$ exists. Hence, the application of this inverse operator to (3.1) yields

$$\begin{cases} \partial_t y(t) = \mathcal{A}(y(t))y(t) + \mathcal{F}(t, y(t)), & t \in J_T, \\ \quad y(0) = y_0, \end{cases} \tag{3.3}$$

where we introduced the mappings

$$\mathcal{A}(\xi) \coloneqq \Lambda(\xi)^{-1}\mathrm{A}, \qquad \mathcal{F}(t,\xi) \coloneqq \Lambda(\xi)^{-1}\mathrm{F}(t,\xi), \qquad t \in J_T,\, \xi \in B_{\mathcal{Y}}(R). \tag{3.4}$$

Furthermore, for $\xi \in B_{\mathcal{Y}}(R)$ we define the state-dependent inner product

$$(\varphi \mid \psi)_{\Lambda(\xi)} \coloneqq (\Lambda(\xi)\varphi \mid \psi)_{\mathcal{X}}, \qquad \varphi, \psi \in \mathcal{X}, \tag{3.5}$$

which is equivalent to $(\cdot \mid \cdot)_{\mathcal{X}}$ due to (3.2). We denote the induced norm by

$$\|\varphi\|_{\Lambda(\xi)}^2 \coloneqq (\varphi \mid \varphi)_{\Lambda(\xi)}, \qquad \varphi \in \mathcal{X}.$$

The discretization presented in the following chapters is based on the wellposedness of (3.3). Despite the absence of a wellposedness result for (3.3) which covers all our examples, we simply assume that the operators from Assumption 3.1 and the initial value $y_0 \in \mathcal{X}$ are chosen such that there exists a unique solution. This is summarized in the following assumption.

Assumption 3.2. *Let Assumption 3.1 be satisfied. The quasilinear Cauchy problem (3.3) has a unique solution y with maximal time of existence $t^*(y_0) > 0$, i.e., for every $T < t^*(y_0)$ the solution y of (3.3) on J_T satisfies*

$$y \in C^1(J_T, \mathcal{X}) \cap C(J_T, B_{\mathcal{Y}}(R)).$$

Given such a solution, the weak formulation of the problem (3.1) considered on $(\mathcal{X}, (\cdot \mid \cdot)_{\mathcal{X}})$ is identical to the weak formulation of (3.3) considered on $(\mathcal{X}, (\cdot \mid \cdot)_{\Lambda(y)})$. Hence, it is sufficient to consider only (3.3) in the following.

The previous assumption is motivated in the subsections by the presentation of wellposedness results for specific examples.

3.2 Example: Westervelt equation

As the first example to illustrate the physical relevance of the general framework presented in the previous section, we consider the Westervelt equation, which is a model in nonlinear acoustics. Based on the Navier Stokes equation and the equation of continuity, it describes the propagation of waves in lossy and compressible fluids, especially for a propagation of multiple wavelengths. For further insight into the physical background and the derivation of the model, we refer to the original work [Westervelt, 1963], as well as [Kaltenbacher, 2015] and [Lerch et al., 2009].

The original version of the Westervelt equation was presented in [Westervelt, 1963]. It states that, for a time interval $J_T = [0, T]$ and a bounded domain $\Omega \subset \mathbb{R}^d$ with $d \in \{1, 2, 3\}$, the acoustic pressure $u : J_T \times \Omega \to \mathbb{R}$ is given as the solution of the quasilinear wave equation

$$(1 - \varkappa u)\partial_t^2 u = c^2 \Delta u + \varkappa(\partial_t u)^2. \tag{3.6}$$

Here, $c > 0$ is the speed of sound and $\varkappa \in \mathbb{R}$ is modeling the nonlinearity of the medium.

Since we consider this problem for given initial values $u_0, v_0 : \Omega \to \mathbb{R}$, we finally obtain

$$\begin{cases} (1 - \varkappa u)\partial_t^2 u = \Delta u + \varkappa(\partial_t u)^2 & \text{on } J_T \times \Omega, \\ \qquad u(0) = u_0, \qquad \partial_t u(0) = v_0 & \text{on } \Omega, \end{cases} \tag{3.7}$$

subject to homogeneous Dirichlet boundary conditions.

Before we review the corresponding wellposedness result, we first write the Westervelt equation as a first order system in order to check that Assumption 3.1 is satisfied. Setting $v = \partial_t u$ in (3.7) implies

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 - \varkappa u \end{pmatrix} \begin{pmatrix} \partial_t u \\ \partial_t v \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \Delta & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ \varkappa v^2 \end{pmatrix}.$$

Hence, (3.1) holds for

$$y = \begin{pmatrix} u \\ v \end{pmatrix}, \quad \Lambda(y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 - \varkappa u \end{pmatrix}, \quad \mathrm{A} = \begin{pmatrix} 0 & 1 \\ \Delta & 0 \end{pmatrix}, \quad \mathrm{F}(y) = \begin{pmatrix} 0 \\ \varkappa v^2 \end{pmatrix}, \quad y_0 = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}. \tag{3.8}$$

We now check Assumption 3.1 for $\mathcal{X} = \mathcal{X}_{\mathcal{V}} \times \mathcal{X}_{\mathcal{H}}$ and $\mathcal{Y} = \mathcal{Y}_{\mathcal{V}} \times \mathcal{Y}_{\mathcal{H}}$, with

$$\mathcal{X}_{\mathcal{V}} = H_0^1(\Omega), \qquad \mathcal{X}_{\mathcal{H}} = L^2(\Omega), \qquad \mathcal{Y}_{\mathcal{V}} = H^2(\Omega) \cap H_0^1(\Omega), \qquad \mathcal{Y}_{\mathcal{H}} = H_0^1(\Omega) \tag{3.9}$$

and the corresponding standard inner products, which are introduced in Chapter 2. Without loss of generality, we assume in the following that the problem is really quasilinear, i.e., we have $\varkappa \neq 0$.

First, the triangle inequality and Sobolev's embedding (2.5) with $\ell = 2$ yield

$$\|1 - \varkappa\varphi\|_{L^\infty(\Omega)} \leq 1 + |\varkappa|\|\varphi\|_{L^\infty(\Omega)} \leq 1 + |\varkappa|C_S R, \qquad \varphi \in B_{\mathcal{Y}_\mathcal{V}}(R),$$

and

$$\|1 - \varkappa\varphi\|_{L^\infty(\Omega)} \geq 1 - |\varkappa|\|\varphi\|_{L^\infty(\Omega)} \geq 1 - |\varkappa|C_S R, \qquad \varphi \in B_{\mathcal{Y}_\mathcal{V}}(R).$$

Thus, for $R < \frac{1}{C_S|\varkappa|}$ (Λ) is satisfied with $c_\Lambda = 1 - |\varkappa|C_S R$ and $C_\Lambda = 1 + |\varkappa|C_S R$.

Furthermore, we have for all $\varphi \in H_0^1(\Omega) \cap H^2(\Omega)$ that $\Delta\varphi \in L^2(\Omega)$. Hence (A) is satisfied with $D(\mathrm{A}) = \mathcal{Y}$.

Finally, we obtain for $v \in B_{\mathcal{Y}_\mathcal{H}}(R)$ the bound

$$\|\varkappa v^2\|_{\mathcal{X}_\mathcal{H}} = |\varkappa|\|v\|^2_{L^4(\Omega)} \leq C|\varkappa|R^2$$

with a constant $C > 0$ depending on Ω and d, where we used the Hölder inequality together with Sobolev's embedding (2.5) in the last step. As this yields (F), Assumption 3.1 is satisfied.

To prove the wellposedness of (3.7), we review the main result of [Dörfler et al., 2016], where the authors consider the problem

$$\partial_t^2(u(t) + K(u(t))) = \Delta u(t), \qquad t \in J_T, \tag{3.10}$$

on a bounded, smooth domain. For the choice $K(u) = -\frac{k}{2}u^2$, this is equivalent to the undamped Westervelt equation. Hence, we get the following wellposedness result corresponding to Assumption 3.2, which follows from [Dörfler et al., 2016, Thm. 4.1].

Theorem 3.3. *Let Ω be a bounded domain with C^3-boundary and R as above. Then, there exists a constant $C_0 \geq 1$ depending on $\varkappa$ and R such that for initial values*

$$u_0 \in \{\varphi \in H^3(\Omega) \cap H_0^1(\Omega) \mid \Delta\varphi|_{\partial\Omega} = 0\}, \qquad v_0 \in H^2(\Omega) \cap H_0^1(\Omega), \tag{3.11a}$$

which satisfy the smallness condition

$$\|(u_0, v_0)\|_\mathcal{Y} < \frac{R}{C_0}, \tag{3.11b}$$

there exists $T > 0$ and a solution u of (3.7) on $J_T = [0, T]$, which satisfies

$$u \in C^2(J_T, H_0^1(\Omega)) \cap C^1(J_T, H^2(\Omega)) \cap C(J_T, H^3(\Omega))$$

and $\|(u(t), \partial_t u(t))\|_\mathcal{Y} < R$ for all $t \leq T$.

Unfortunately, we have to refine the space $\mathcal{Y}$ for the case $d \in \{2, 3\}$ in order to ensure $\mathcal{Y} \hookrightarrow C(\Omega)^2$, which is essential for the discretization in Section 5.1. In particular, we have to assume for $d \in \{2, 3\}$ that the statement of Theorem 3.3 is also true for $\mathcal{Y} = \mathcal{Y}_\mathcal{V} \times \mathcal{Y}_\mathcal{H}$ with

$$\mathcal{Y}_\mathcal{V} = H^2(\Omega) \cap H_0^1(\Omega), \qquad \mathcal{Y}_\mathcal{H} = H^2(\Omega) \cap H_0^1(\Omega). \tag{3.12}$$

We point out that Assumption 3.1 is still satisfied.

Moreover, even with this refinement the regularity of the solution in the previous result is not sufficient for the error estimates for the Westervelt equation in Section 5.1 and Section 8.1. Instead, these estimates are based on the additional assumption $(u, \partial_t u) \in C^1(J_T, \mathcal{Z})$, with $\mathcal{Z} = \mathcal{Z}_\mathcal{V} \times \mathcal{Z}_\mathcal{H}$ given by

$$\mathcal{Z}_\mathcal{V} = H^{p+1}(\Omega) \cap H_0^1(\Omega), \qquad \mathcal{Z}_\mathcal{H} = H^p(\Omega) \cap H_0^1(\Omega), \tag{3.13}$$

for some $p \in \mathbb{N}$.

To conclude this section, we state the following remark on the ambiguity of the term "Westervelt equation".

Remark 3.4. *In the literature the term "Westervelt equation" does not only refer to the original model (3.6), but also to refined models with strong damping. Especially in the engineering literature, the term Westervelt equation is also used for the model*

$$(1 - \varkappa u)\partial_t^2 u = c^2 \Delta u + bc^{-2}\partial_t^3 u + \varkappa(\partial_t u)^2, \tag{3.14}$$

where $b > 0$ is the sound diffusivity, cf. [Kaltenbacher, 2015] and [Lerch et al., 2009]. Based on the approximation $\partial_t^2 u \approx c^2 \Delta u$, (3.14) further implies the following strongly damped version of the Westervelt equation, i.e.,

$$(1 - \varkappa u)\partial_t^2 u = c^2 \Delta u + b \Delta \partial_t u + \varkappa(\partial_t u)^2, \tag{3.15}$$

cf. [Kaltenbacher and Lasiecka, 2009]. Note that the wellposedness of (3.15) is considered in [Kaltenbacher and Lasiecka, 2009] and [Meyer and Wilke, 2011]. As the problem shows a rather parabolic than hyperbolic behavior due to the strong daming, the authors even obtain exponential decay of the energy for small data, which finally yields global existence.

In the next section we introduce another specific example, the quasilinear Maxwell equations.

3.3 Example: Maxwell equations

As a second example we consider quasilinear Maxwell equations with Kerr nonlinearity. To this end, we briefly introduce Maxwell equations and the material laws describing the nonlinearity. Next, we discuss the wellposedness of this system based on the analysis presented in [Spitz, 2019] for a general class of quasilinear Maxwell equations.

For a finite time interval $J_T = [0, T]$ and a bounded domain $\Omega \subset \mathbb{R}^3$, we denote the magnetic field by $\mathcal{H} : J_T \times \Omega \to \mathbb{R}^3$ and the electric field by $\mathcal{E} : J_T \times \Omega \to \mathbb{R}^3$. Further, $\mathcal{B} : J_T \times \Omega \to \mathbb{R}^3$ is the magnetic induction and $\mathcal{D} : J_T \times \Omega \to \mathbb{R}^3$ is the electric displacement. With the electric current density $\mathcal{J} : J_T \times \Omega \to \mathbb{R}^3$ and the electric charge density $\rho : J_T \times \Omega \to \mathbb{R}$, we end up with the macroscopic Maxwell equations in differential form

$$\partial_t \mathcal{B}(t, x) = -\nabla\times \mathcal{E}(t, x), \qquad t \in J_T,\ x \in \Omega, \tag{3.16a}$$

$$\partial_t \mathcal{D}(t, x) = \nabla\times \mathcal{H}(t, x) - \mathcal{J}(t, x), \qquad t \in J_T,\ x \in \Omega, \tag{3.16b}$$

$$\nabla \cdot \mathcal{B}(t, x) = 0, \qquad t \in J_T,\ x \in \Omega, \tag{3.16c}$$

$$\nabla \cdot \mathcal{D}(t, x) = \rho(t, x), \qquad t \in J_T,\ x \in \Omega, \tag{3.16d}$$

where the differential operators $\nabla\times$ and $\nabla\cdot$ denote the curl and divergence, respectively.

We further differentiate (3.16d) with respect to time and use (3.16b) together with the identity $\nabla \cdot (\nabla\times f) = 0$, which holds for all sufficiently smooth functions $f : \mathbb{R}^3 \to \mathbb{R}^3$. This implies the continuity equation

$$\partial_t \rho(t, x) = -\nabla \cdot \mathcal{J}(t, x), \qquad t \in J_T,\ x \in \Omega, \tag{3.17}$$

which is a coupling condition for the current density and the charge density.

Based on the continuity equation (3.17), we get that the equations (3.16a) and (3.16b) are sufficient to describe the time evolution of the system, as the equations (3.16c) and (3.16d) are time invariant. More precisely, if these equations are satisfied at the initial time, i.e.,

$$\nabla \cdot \mathcal{B}(0, x) = 0, \qquad \nabla \cdot \mathcal{D}(0, x) = \rho(0, x), \qquad x \in \Omega, \tag{3.18}$$

they stay true as long as (3.16a), (3.16b) and (3.17) hold true. This follows again from differentiating the divergence equations with respect to time and using the curl equations.

Hence, if (3.18) is satisfied, we are left with at most seven independent equations to describe the time evolution of 16 unknowns $\mathcal{B}$, $\mathcal{D}$, $\mathcal{H}$, $\mathcal{E}$, $\mathcal{J}$, and ρ. Therefore, we need further relations to get a wellposed system, i.e., we impose in the following the so-called constitutive equations, which relate the magnetic induction $\mathcal{B}$ and the electric displacement $\mathcal{D}$ to the magnetic and electric field $\mathcal{H}$ and $\mathcal{E}$, respectively. We further assume a constitutive relation for the coupling of the electric current density $\mathcal{J}$ and the electric field $\mathcal{E}$. In particular, we assume the existence of mappings $\theta_{\mathcal{H}}, \theta_{\mathcal{E}} : \mathbb{R}^6 \to \mathbb{R}^3$ and $\sigma_{\mathcal{E}} : \mathbb{R}^6 \to \mathbb{R}^{3\times 3}$, with

$$\mathcal{B} = \theta_{\mathcal{H}}(\mathcal{H}, \mathcal{E}), \qquad \mathcal{D} = \theta_{\mathcal{E}}(\mathcal{H}, \mathcal{E}), \qquad \mathcal{J} = \sigma_{\mathcal{E}}(\mathcal{H}, \mathcal{E})\mathcal{E}, \qquad \text{on } J_T \times \Omega.$$

Inserting these relations into (3.16) finally yields

$$\begin{aligned} \partial_t \theta_{\mathcal{H}}(\mathcal{H}, \mathcal{E}) &= -\nabla\times \mathcal{E}, && \text{on } J_T \times \Omega, \\ \partial_t \theta_{\mathcal{E}}(\mathcal{H}, \mathcal{E}) &= \nabla\times \mathcal{H} - \sigma_{\mathcal{E}}(\mathcal{H}, \mathcal{E})\mathcal{E}, && \text{on } J_T \times \Omega, \end{aligned} \tag{3.19}$$

which is a coupled system for the magnetic and electric field. The electric charge density is then given by

$$\partial_t \rho = -\nabla \cdot (\sigma_{\mathcal{E}}(\mathcal{H}, \mathcal{E})\mathcal{E}), \qquad \text{on } J_T \times \Omega.$$

Finally, in order to get a wellposed system, we consider (3.19) subject to initial conditions

$$\mathcal{H}(0) = \mathcal{H}_0, \qquad \mathcal{E}(0) = \mathcal{E}_0, \qquad \text{on } \Omega,$$

for $\mathcal{H}_0, \mathcal{E}_0 : \Omega \to \mathbb{R}^3$, and homogeneous perfectly conducting boundary conditions

$$\mathcal{E} \times \nu = 0, \qquad \mathcal{B} \cdot \nu = 0, \qquad \text{on } \partial\Omega, \tag{3.20}$$

with the outer unit normal vector ν of Ω. As shown in [Spitz, 2017, Lem. 7.25] under suitable assumptions on $\theta_{\mathcal{H}}$, $\theta_{\mathcal{E}}$ and $\sigma_{\mathcal{E}}$, this condition is again time invariant, i.e., it is sufficient to impose $\mathcal{B}(0) \cdot \nu = 0$ or equivalently $\theta_{\mathcal{H}}(\mathcal{H}(0), \mathcal{E}(0)) \cdot \nu = 0$.

As before, differentiating the boundary condition for $\mathcal{B}$ with respect to time and using (3.16a) yields that this condition is time invariant, i.e., it is sufficient to impose $\mathcal{B}(0)\cdot\nu = 0$ or equivalently $\theta_{\mathcal{H}}(\mathcal{H}(0), \mathcal{E}(0)) \cdot \nu = 0$.

Based on the mappings

$$\Lambda(y) = \begin{pmatrix} \partial_{\mathcal{H}}\theta_{\mathcal{H}}(\mathcal{H},\mathcal{E}) & \partial_{\mathcal{E}}\theta_{\mathcal{H}}(\mathcal{H},\mathcal{E}) \\ \partial_{\mathcal{H}}\theta_{\mathcal{E}}(\mathcal{H},\mathcal{E}) & \partial_{\mathcal{E}}\theta_{\mathcal{E}}(\mathcal{H},\mathcal{E}) \end{pmatrix}, \quad \mathrm{A} = \begin{pmatrix} 0 & -\nabla\times \\ \nabla\times & 0 \end{pmatrix}, \quad \mathrm{F}(y) = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_{\mathcal{E}}(\mathcal{H},\mathcal{E}) \end{pmatrix} \tag{3.21}$$

for $y = (\mathcal{H}, \mathcal{E})$, we see that the quasilinear Maxwell equations (3.19) fit in the general framework (3.1) with initial value $y_0 = (\mathcal{H}_0, \mathcal{E}_0)$.

Although the procedure presented in the following chapters is, under suitable assumptions on the constitutive relations, also applicable to more general problems, we focus in the following on an instantaneous Kerr-type nonlinearity in an isotropic medium. Namely, for the nonlinear susceptibility $\chi \in L^\infty(\Omega)$, we set

$$\theta_{\mathcal{H}}(\mathcal{H},\mathcal{E}) = \mathcal{H}, \qquad \theta_{\mathcal{E}}(\mathcal{H},\mathcal{E}) = (1 + \chi|\mathcal{E}|^2)\mathcal{E}.$$

For further information on the derivation of these relations, see for instance [Busch et al., 2007] and [Pototschnig et al., 2009]. Moreover, we set $\sigma_{\mathcal{E}} \equiv 0$ for the sake of presentation. Based on (3.21), we thus obtain

$$\Lambda(y) = \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & (1+\chi|\mathcal{E}|^2)\,\mathrm{Id} + 2\chi(\mathcal{E}\otimes\mathcal{E}) \end{pmatrix}, \quad \mathrm{A} = \begin{pmatrix} 0 & -\nabla\times \\ \nabla\times & 0 \end{pmatrix}, \quad \mathrm{F}(y) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \tag{3.22}$$

where $\otimes$ denotes the Kronecker product. Finally, this yields Maxwell equations with Kerr nonlinearities, which are given by

$$\begin{cases} \partial_t\mathcal{H} = -\nabla\times\mathcal{E}, & \text{on } J_T\times\Omega, \\ ((1+\chi|\mathcal{E}|^2)\,\mathrm{Id} + 2\chi(\mathcal{E}\otimes\mathcal{E}))\partial_t\mathcal{E} = \nabla\times\mathcal{H}, & \text{on } J_T\times\Omega, \\ \mathcal{H}(0) = \mathcal{H}_0, \qquad \mathcal{E}(0) = \mathcal{E}_0 & \text{on } \Omega, \end{cases} \tag{3.23}$$

subject to homogeneous perfectly conducting boundary conditions.

We now introduce spaces $\mathcal{X} = \mathcal{X}_{\mathcal{V}} \times \mathcal{X}_{\mathcal{H}}$ and $\mathcal{Y} = \mathcal{Y}_{\mathcal{V}} \times \mathcal{Y}_{\mathcal{H}}$, with

$$\mathcal{X}_{\mathcal{V}} = L^2(\Omega)^3, \quad \mathcal{X}_{\mathcal{H}} = L^2(\Omega)^3, \quad \mathcal{Y}_{\mathcal{V}} = H^2(\Omega)^3, \quad \mathcal{Y}_{\mathcal{H}} = \{\varphi \in H^2(\Omega)^3 \mid \varphi\times\nu = 0\},$$

where ν denotes the outer unit normal vector of Ω. All spaces are equipped with their standard inner products. We emphasize that one has to be cautious here not to confuse the notation of $\mathcal{H}$ and $\mathcal{H}$. However, we stick to this notation since it is consistent with the abstract framework and the standard notation for the Maxwell equations.

In order to check Assumption 3.1, we assume without loss of generality that the problem is truly quasilinear, i.e., we have $\|\chi\|_{L^\infty(\Omega)} > 0$.

First of all, (F) is trivially satisfied. Furthermore, we obtain for

$$D(\mathrm{A}) = H(\mathrm{curl},\Omega) \times H_0(\mathrm{curl},\Omega)$$

from [Monk, 2003, Thm. 3.33] the relation $\mathcal{Y} \subset D(\mathrm{A})$. As $H(\mathrm{curl},\Omega)$ and $H_0(\mathrm{curl},\Omega)$ correspond to the closure of $C^\infty(\overline{\Omega})^3$ and $C_0^\infty(\Omega)^3$ with repect to the norm of $H(\mathrm{curl},\Omega)$, respectively, (A) is satisfied.

Finally, we focus on (Λ). The triangle inequality and (2.2) for $\xi \in \mathcal{Y}_{\mathcal{V}}$ and $\varphi \in \mathcal{X}_{\mathcal{H}}$ yield

$$\big\|\chi(x)|\xi(x)|^2\varphi(x) + 2\chi(x)(\xi(x)\otimes\xi(x))\varphi(x)\big\|_2 \le 9|\chi(x)|\|\xi(x)\|_\infty^2\|\varphi(x)\|_2, \qquad x\in\Omega.$$

From the definition (2.4) of norms for vector-valued functions and Sobolev's embedding (2.5), we hence obtain

$$\begin{aligned}\|(1+\chi|\xi|^2+2\chi(\xi\otimes\xi))\varphi\|_{\mathcal{X}_{\mathcal{H}}} &= \Big\|\|(1+\chi|\xi|^2+2\chi(\xi\otimes\xi))\varphi\|_2\Big\|_{L^2(\Omega)}\\ &\le (1+9\|\chi\|_{L^\infty(\Omega)}C_{\mathrm{S}}^2R^2)\|\varphi\|_{\mathcal{X}_{\mathcal{H}}}.\end{aligned}$$

Thus, for $R < \big(3\|\chi\|_{L^\infty(\Omega)}^{\frac{1}{2}}C_{\mathrm{S}}\big)^{-1}$ and $C_\Lambda = 1+9\|\chi\|_{L^\infty(\Omega)}C_{\mathrm{S}}^2R^2$, this yields the upper bound in (3.2). As the lower bound can be proven similarly with $c_\Lambda = 1-9\|\chi\|_{L^\infty(\Omega)}C_{\mathrm{S}}^2R^2$, we have verified Assumption 3.1.

Moreover, we emphasize that the upper bound for the radius R is not necessary if the nonlinear susceptibility χ is non-negative. In this case, (Λ) and thus also Assumption 3.1 is satisfied with $R>0$ arbitrary.

For the wellposedness of this problem, we rely on [Spitz, 2019, Thm. 5.3], where a very general class of quasilinear Maxwell equations is analyzed. Before stating the theorem, we briefly comment on compatibility conditions, which are essential for this result.

The compatibility conditions of order $m \in \mathbb{N}$ state that the electric field and its derivatives with respect to time satisfy the boundary condition (3.20) at the initial time $t=0$, i.e., we have

$$\nu \times \partial_t^p \mathcal{E}(0) = 0, \qquad\qquad p = 0,\dots,m-1.$$

Using the Maxwell equations (3.23), this can be traced back to an assumption on the initial values and the nonlinear susceptibility χ. For further details on these conditions, we refer to [Spitz, 2019, Sec. 2].

Theorem 3.5. *For $m\in\mathbb{N}$, let $\Omega\subset\mathbb{R}^3$ be a bounded domain with $C^{\max\{m,3\}+2}$-boundary. Further, let $\chi\in C^m(\Omega)$ and R as above. For all $\alpha\in\mathbb{N}_0^3$ with $|\alpha|\le m$, we assume $\|\partial^\alpha\chi\|_{L^\infty(\Omega)}<\infty$. Then, if the initial values $y_0=(\mathcal{H}_0,\mathcal{E}_0)$ satisfy*

$$y_0 \in H^m(\Omega)^6 \cap B_{\mathcal{Y}}(R),$$

as well as the compatibility conditions of order m, the Maxwell equations with Kerr nonlinearity (3.23) have a unique solution with maximal time of existence $t^*(y_0)>0$, i.e., for all $T<t^*(y_0)$, there exists a unique solution $y=(\mathcal{H},\mathcal{E})$ of (3.23) on $J_T=[0,T]$, which satisfies

$$y \in \bigcap_{j=0}^{m} C^j\big((0,t^*(y_0)),H^{m-j}(\Omega)^6\big)$$

and $\|y(t)\|_{\mathcal{Y}}<R$ for all $t\le T$.

As for the Westervelt equation, we require additional regularity of the solution for the error estimates in Section 5.2 and Section 8.2, i.e., we need $y\in C^1(J_T,\mathcal{Z})$, where $\mathcal{Z}_{\mathcal{V}}\times\mathcal{Z}_{\mathcal{H}}$ is for some $p\in\mathbb{N}$ given by

$$\mathcal{Z}_{\mathcal{V}} = H^{p+1}(\Omega)^3, \qquad\qquad \mathcal{Z}_{\mathcal{H}} = H^{p+1}(\Omega)^3. \tag{3.24}$$

However, we stress that with Theorem 3.5, this can be traced back to assumptions on the data.

To conclude this section, we emphasize that, for the sake of presentation, we consider a rather simple model problem here. Nevertheless, note that both the wellposedness result from [Spitz, 2019, Thm. 5.3] as well as the abstract analysis presented in the following are also suitable for more general problems, e.g., including different constitutive relations or non-trivial boundary conditions.

CHAPTER 4

Space discretization of abstract problems

In this chapter we investigate the space discretization of quasilinear wave-type problems of the form (3.1). Thus, we introduce in Section 4.1 the abstract discrete framework. In Section 4.2 we present a brief excursion to semigroups for nonautonomous Cauchy problems, which is fundamental for the error analysis of the space discretization in Section 4.3. Moreover, we present a refined version of the error estimate in Section 4.4, based on further assumptions on the nonlinearities.

4.1 General setting

For the space discretization of quasilinear evolution equations, we employ a finite-dimensional vector space $\boldsymbol{\mathcal{V}}$ in which we seek the approximation $\boldsymbol{y}$ of the exact solution $y \in \mathcal{X}$ of (3.3). If this space is equipped with the inner product $(\cdot \mid \cdot)_{\boldsymbol{\mathcal{X}}}$, which corresponds to the inner product of $\mathcal{X}$ we introduce the simplifying notation $\boldsymbol{\mathcal{X}} = (\boldsymbol{\mathcal{V}}, (\cdot \mid \cdot)_{\boldsymbol{\mathcal{X}}})$. Furthermore, we introduce the normed vector space $\boldsymbol{\mathcal{Y}} = (\boldsymbol{\mathcal{V}}, \|\cdot\|_{\boldsymbol{\mathcal{Y}}})$, where $\|\cdot\|_{\boldsymbol{\mathcal{Y}}}$ corresponds to the norm induced by the inner product of $\mathcal{Y}$. Finally, we denote by $h > 0$ the discretization parameter; e.g., for the discretization with finite elements, this corresponds to the maximal diameter of the mesh elements.

As the wellposedness of quasilinear wave-type equations in many cases depends on the smoothness of the boundary of the domain, which will not necessarily be available in the discrete setting, it is only natural to consider nonconforming space discretizations here, i.e., we allow for $\boldsymbol{\mathcal{X}} \not\subset \mathcal{X}$.

As $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$ are finite-dimensional spaces, all norms are equivalent with constants depending on h, i.e., we have

$$\frac{1}{\boldsymbol{C}_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}}(h)} \|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}} \leq \|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{Y}}} \leq \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h) \|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}}, \qquad \boldsymbol{\xi} \in \boldsymbol{\mathcal{Y}}. \tag{4.1}$$

Note that, as $\boldsymbol{\mathcal{X}}$ is the weaker space compared to $\boldsymbol{\mathcal{Y}}$, the dependency of $\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)$ on the space discretization parameter is really mandatory. Hence, the first bound is called inverse estimate. However, there are also examples where even $\boldsymbol{C}_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}}(h)$ depends on h, cf. Section 5.1.

With respect to the specific examples in Chapter 3, we obtain for both the Westervelt equation and the Maxwell equations $\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h) \sim h^{-\frac{d}{2}}$, where $d \in \mathbb{N}$ denotes the spatial dimension. Moreover, for the Westervelt equation we have $\boldsymbol{C}_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}}(h) \sim h^{-1}$, whereas for the Maxwell equations $\boldsymbol{C}_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}}(h)$ is independent of h.

With discretizations $\mathbf{\Lambda}$, $\mathbf{A}$, and $\mathbf{F}$ of Λ, A, and F, respectively, which are specified in the following assumptions, we obtain the following discrete system

$$\begin{cases} \mathbf{\Lambda}(\boldsymbol{y}(t))\partial_t \boldsymbol{y}(t) = \mathbf{A}\boldsymbol{y}(t) + \mathbf{F}(t, \boldsymbol{y}(t)), & t \in J_T, \\ \qquad \boldsymbol{y}(0) = \boldsymbol{y}_0. \end{cases} \tag{4.2}$$

The analysis below is based on the following properties of the discrete operators.

Assumption 4.1. *There exists $\boldsymbol{R} > 0$ such that the discrete operators satisfy the following properties uniformly in $h > 0$.*

($\mathbf{\Lambda}$) $\{\mathbf{\Lambda}(\boldsymbol{\xi}) \mid \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})\} \subset \mathcal{L}(\boldsymbol{\mathcal{X}})$ *is a family of symmetric operators, which are uniformly positive definite and bounded, i.e., there are constants $\boldsymbol{c}_{\mathbf{\Lambda}}, \boldsymbol{C}_{\mathbf{\Lambda}} > 0$ such that*

$$\boldsymbol{c}_{\mathbf{\Lambda}} \|\varphi\|_{\boldsymbol{\mathcal{X}}}^2 \le (\mathbf{\Lambda}(\boldsymbol{\xi})\varphi \mid \varphi)_{\boldsymbol{\mathcal{X}}}, \qquad \|\mathbf{\Lambda}(\boldsymbol{\xi})\|_{\mathcal{L}(\boldsymbol{\mathcal{X}})} \le \boldsymbol{C}_{\mathbf{\Lambda}}, \qquad \varphi \in \boldsymbol{\mathcal{X}},\ \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{4.3}$$

holds. Furthermore, there are constants $\boldsymbol{L}_{\mathbf{\Lambda}}^{\boldsymbol{\mathcal{X}}}, \boldsymbol{L}_{\mathbf{\Lambda}}^{\boldsymbol{\mathcal{Y}}} > 0$ such that

$$\|\mathbf{\Lambda}(\boldsymbol{\varphi}) - \mathbf{\Lambda}(\boldsymbol{\psi})\|_{\mathcal{L}(\boldsymbol{\mathcal{X}})} \le \boldsymbol{L}_{\mathbf{\Lambda}}^{\boldsymbol{\mathcal{X}}} \|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_{\boldsymbol{\mathcal{Y}}}, \qquad \boldsymbol{\varphi}, \boldsymbol{\psi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}), \tag{4.4a}$$

$$\|\mathbf{\Lambda}(\boldsymbol{\varphi}) - \mathbf{\Lambda}(\boldsymbol{\psi})\|_{\mathcal{L}(\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}})} \le \boldsymbol{L}_{\mathbf{\Lambda}}^{\boldsymbol{\mathcal{Y}}} \|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}}, \qquad \boldsymbol{\varphi}, \boldsymbol{\psi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{4.4b}$$

hold.

($\mathbf{A}$) $\mathbf{A} : \boldsymbol{\mathcal{X}} \to \boldsymbol{\mathcal{X}}$ *is dissipative in $\boldsymbol{\mathcal{X}}$, i.e.,*

$$(\mathbf{A}\boldsymbol{\xi} \mid \boldsymbol{\xi})_{\boldsymbol{\mathcal{X}}} \le 0, \qquad \boldsymbol{\xi} \in \boldsymbol{\mathcal{X}} \tag{4.5}$$

holds.

($\mathbf{F}$) *We have $\mathbf{F} : J_T \times B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \to \boldsymbol{\mathcal{X}}$, which is continuous in time and bounded in $\boldsymbol{\mathcal{Y}}$, i.e., there is a constant $\boldsymbol{C}_{\mathbf{F}} > 0$ such that*

$$\|\mathbf{F}(t, \boldsymbol{\xi})\|_{\boldsymbol{\mathcal{Y}}} \le \boldsymbol{C}_{\mathbf{F}}, \qquad t \in J_T,\ \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{4.6}$$

holds. Let further $\mathbf{F}$ be Lipschitz continuous in the second argument, i.e., there is a constant $\boldsymbol{L}_{\mathbf{F}} > 0$ such that

$$\|\mathbf{F}(t, \boldsymbol{\varphi}) - \mathbf{F}(t, \boldsymbol{\psi})\|_{\boldsymbol{\mathcal{X}}} \le \boldsymbol{L}_{\mathbf{F}} \|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}}, \qquad t \in J_T,\ \boldsymbol{\varphi}, \boldsymbol{\psi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{4.7}$$

holds.

In the following, we always assume that $\boldsymbol{R} > 0$ is chosen such that Assumption 4.1 is satisfied.

As in the continuous case, these assumptions yield for $\boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$ the discrete state-dependent inner product

$$(\boldsymbol{\varphi} \mid \boldsymbol{\psi})_{\mathbf{\Lambda}(\boldsymbol{\xi})} := (\mathbf{\Lambda}(\boldsymbol{\xi})\boldsymbol{\varphi} \mid \boldsymbol{\psi})_{\boldsymbol{\mathcal{X}}}, \qquad \boldsymbol{\varphi}, \boldsymbol{\psi} \in \boldsymbol{\mathcal{X}}. \tag{4.8}$$

We again denote the induced norm by

$$\|\boldsymbol{\varphi}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\xi})} \coloneqq (\boldsymbol{\varphi} \mid \boldsymbol{\varphi})_{\boldsymbol{\Lambda}(\boldsymbol{\xi})}, \qquad \boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}}. \tag{4.9}$$

In the next lemma, we state some properties of the state-dependent inner product and its induced norm.

Lemma 4.2. *For $\boldsymbol{\zeta} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$, the norm equivalence*

$$\boldsymbol{c}_{\boldsymbol{\Lambda}}\|\boldsymbol{\xi}\|^2_{\boldsymbol{\mathcal{X}}} \leq \|\boldsymbol{\xi}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\zeta})} \leq \boldsymbol{C}_{\boldsymbol{\Lambda}}\|\boldsymbol{\xi}\|^2_{\boldsymbol{\mathcal{X}}}, \qquad \boldsymbol{\xi} \in \boldsymbol{\mathcal{X}} \tag{4.10}$$

holds. Let further

$$\boldsymbol{z} \in C^1(J_T, \boldsymbol{\mathcal{Y}}) \cap C(J_T, B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})),$$

with $\|\partial_t \boldsymbol{z}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}^{\partial_t}$ for some $\boldsymbol{R}^{\partial_t} > 0$. Then, the state-dependent norm depends continuously on time in the sense that the estimate

$$\|\boldsymbol{\xi}\|_{\boldsymbol{\Lambda}(\boldsymbol{z}(t))} \leq (1 + C'|t-s|)\|\boldsymbol{\xi}\|_{\boldsymbol{\Lambda}(\boldsymbol{z}(s))} \leq \mathrm{e}^{C'|t-s|}\|\boldsymbol{\xi}\|_{\boldsymbol{\Lambda}(\boldsymbol{z}(s))}, \qquad s,t \in J_T,\ \boldsymbol{\xi} \in \boldsymbol{\mathcal{X}} \tag{4.11}$$

is satisfied with the constant $C' = \frac{1}{2}\boldsymbol{L}^{\boldsymbol{\mathcal{X}}}_{\boldsymbol{\Lambda}}\boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}\boldsymbol{R}^{\partial_t}$.

Proof. The norm equivalence (4.10) is a direct consequence of (4.3). To prove (4.11), let $\boldsymbol{z} \in C^1(J_T, B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}^{\partial_t})) \cap C(J_T, B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}))$ and $\boldsymbol{\xi} \in \boldsymbol{\mathcal{X}}$. Without loss of generality, let further $s, t \in J_T$ with $s \leq t$. This yields

$$\begin{aligned}
\|\boldsymbol{\xi}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{z}(t))} &= (\boldsymbol{\Lambda}(\boldsymbol{z}(t))\boldsymbol{\xi} \mid \boldsymbol{\xi})_{\boldsymbol{\mathcal{X}}} \\
&= ((\boldsymbol{\Lambda}(\boldsymbol{z}(t)) - \boldsymbol{\Lambda}(\boldsymbol{z}(s)))\boldsymbol{\xi} \mid \boldsymbol{\xi})_{\boldsymbol{\mathcal{X}}} + (\boldsymbol{\Lambda}(\boldsymbol{z}(s))\boldsymbol{\xi} \mid \boldsymbol{\xi})_{\boldsymbol{\mathcal{X}}} \\
&\leq \boldsymbol{L}^{\boldsymbol{\mathcal{X}}}_{\boldsymbol{\Lambda}}\|\boldsymbol{z}(t) - \boldsymbol{z}(s)\|_{\boldsymbol{\mathcal{Y}}}\|\boldsymbol{\xi}\|^2_{\boldsymbol{\mathcal{X}}} + \|\boldsymbol{\xi}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{z}(s))} \\
&\leq \Big(\boldsymbol{L}^{\boldsymbol{\mathcal{X}}}_{\boldsymbol{\Lambda}}\boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}\int_s^t \|\partial_t \boldsymbol{z}(r)\|_{\boldsymbol{\mathcal{Y}}}\,\mathrm{d}r + 1\Big)\|\boldsymbol{\xi}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{z}(s))},
\end{aligned}$$

where we first used the Lipschitz continuity (4.4a) of $\boldsymbol{\Lambda}$. For the last inequality, we applied the fundamental theorem of calculus together with the norm equivalence (4.10). Finally, we deduce from $\partial_t \boldsymbol{z}(r) \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}^{\partial_t})$ the inequality

$$\begin{aligned}
\|\boldsymbol{\xi}\|_{\boldsymbol{\Lambda}(\boldsymbol{z}(t))} &\leq \Big(\boldsymbol{L}^{\boldsymbol{\mathcal{X}}}_{\boldsymbol{\Lambda}}\boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}\boldsymbol{R}^{\partial_t}|t-s| + 1\Big)^{\frac{1}{2}}\|\boldsymbol{\xi}\|_{\boldsymbol{\Lambda}(\boldsymbol{z}(s))} \\
&\leq \Big(1 + \frac{1}{2}\boldsymbol{L}^{\boldsymbol{\mathcal{X}}}_{\boldsymbol{\Lambda}}\boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}\boldsymbol{R}^{\partial_t}|t-s|\Big)\|\boldsymbol{\xi}\|_{\boldsymbol{\Lambda}(\boldsymbol{z}(s))} \\
&\leq \mathrm{e}^{\frac{1}{2}\boldsymbol{L}^{\boldsymbol{\mathcal{X}}}_{\boldsymbol{\Lambda}}\boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}\boldsymbol{R}^{\partial_t}|t-s|}\|\boldsymbol{\xi}\|_{\boldsymbol{\Lambda}(\boldsymbol{z}(s))},
\end{aligned}$$

which completes the proof. □

As in the continuous case, we observe that under assumption ($\boldsymbol{\Lambda}$), the family of discrete inverse operators $\{\boldsymbol{\Lambda}(\boldsymbol{\xi})^{-1} \mid \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})\} \subset \mathcal{L}(\boldsymbol{\mathcal{X}})$ is well defined. Application of the discrete inverse operator yields that (4.2) is equivalent to

$$\begin{cases} \partial_t \boldsymbol{y}(t) = \boldsymbol{\mathcal{A}}(\boldsymbol{y}(t))\boldsymbol{y}(t) + \boldsymbol{\mathcal{F}}(t, \boldsymbol{y}(t)), & t \in J_T, \\ \boldsymbol{y}(0) = \boldsymbol{y}_0, \end{cases} \tag{4.12}$$

with the mappings

$$\boldsymbol{\mathcal{A}}(\boldsymbol{\xi}) := \boldsymbol{\Lambda}(\boldsymbol{\xi})^{-1}\mathbf{A}, \qquad \boldsymbol{\mathcal{F}}(t,\boldsymbol{\xi}) := \boldsymbol{\Lambda}(\boldsymbol{\xi})^{-1}\mathbf{F}(t,\boldsymbol{\xi}), \qquad t \in J_T,\ \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}), \tag{4.13}$$

cf. (3.3). Similarly to the continuous case, we get that if the solution $\boldsymbol{y}$ of (4.12) satisfies

$$\boldsymbol{y} \in C^1(J_T, \boldsymbol{\mathcal{X}}) \cap C(J_T, B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})),$$

then the weak form of (4.2) considered on $(\boldsymbol{\mathcal{X}}, (\cdot \mid \cdot)_{\boldsymbol{\mathcal{X}}})$ is identical to the weak form of (4.12) considered on $(\boldsymbol{\mathcal{X}}, (\cdot \mid \cdot)_{\boldsymbol{\Lambda}(\boldsymbol{y})})$.

Based on Assumption 4.1, which states the properties of the discrete operators appearing in (4.2), we show in the following lemma that the operators appearing in (4.12) are again Lipschitz continuous.

Lemma 4.3. *There are constants $\boldsymbol{L}_{\boldsymbol{\mathcal{A}}}, \boldsymbol{L}_{\boldsymbol{\mathcal{F}}} > 0$ such that for all $\boldsymbol{\varphi}, \boldsymbol{\psi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$, the bounds*

$$\|(\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}) - \boldsymbol{\mathcal{A}}(\boldsymbol{\psi}))\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}} \le \boldsymbol{L}_{\boldsymbol{\mathcal{A}}} \|\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi})\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{Y}}} \|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}}, \qquad \boldsymbol{\xi} \in \boldsymbol{\mathcal{X}}, \tag{4.14}$$

$$\|\boldsymbol{\mathcal{F}}(t,\boldsymbol{\varphi}) - \boldsymbol{\mathcal{F}}(t,\boldsymbol{\psi})\|_{\boldsymbol{\mathcal{X}}} \le \boldsymbol{L}_{\boldsymbol{\mathcal{F}}} \|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}}, \qquad t \in J_T, \tag{4.15}$$

hold.

Proof. Let $\boldsymbol{\xi} \in \boldsymbol{\mathcal{X}}$ and $\boldsymbol{\varphi}, \boldsymbol{\psi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$ be chosen arbitrarily. First, we obtain from (4.3) the bound

$$\|\boldsymbol{\Lambda}(\xi)^{-1}\|_{\mathcal{L}(\boldsymbol{\mathcal{X}})} \le \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}. \tag{4.16}$$

To prove (4.14), we hence deduce from (4.13) and the Lipschitz continuity (4.4b) of $\boldsymbol{\Lambda}$

$$\begin{aligned}
\|(\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}) - \boldsymbol{\mathcal{A}}(\boldsymbol{\psi}))\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}} &\le \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1} \|\boldsymbol{\Lambda}(\boldsymbol{\psi})(\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}) - \boldsymbol{\mathcal{A}}(\boldsymbol{\psi}))\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}} \\
&\le \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1} \|(\boldsymbol{\Lambda}(\boldsymbol{\psi})\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}) - \mathbf{A})\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}} \\
&\le \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1} \|(\boldsymbol{\Lambda}(\boldsymbol{\psi}) - \boldsymbol{\Lambda}(\boldsymbol{\varphi}))\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi})\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}} \\
&\le \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1} \boldsymbol{L}_{\boldsymbol{\Lambda}}^{\boldsymbol{\mathcal{Y}}} \|\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi})\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{Y}}} \|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}}.
\end{aligned}$$

To derive inequality (4.15), let $t \in J_T$ be arbitrary. We then obtain from (4.13) and the Lipschitz continuity of both $\boldsymbol{\Lambda}$ and $\mathbf{F}$ from (4.4b) and (4.7), respectively, the bound

$$\begin{aligned}
\|\boldsymbol{\mathcal{F}}(t,\boldsymbol{\varphi}) - \boldsymbol{\mathcal{F}}(t,\boldsymbol{\psi})\|_{\boldsymbol{\mathcal{X}}} &\le \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1} \|\boldsymbol{\Lambda}(\boldsymbol{\psi})(\boldsymbol{\mathcal{F}}(t,\boldsymbol{\varphi}) - \boldsymbol{\mathcal{F}}(t,\boldsymbol{\psi}))\|_{\boldsymbol{\mathcal{X}}} \\
&\le \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1} \|(\boldsymbol{\Lambda}(\boldsymbol{\psi}) - \boldsymbol{\Lambda}(\boldsymbol{\varphi}))\boldsymbol{\mathcal{F}}(t,\boldsymbol{\varphi})\|_{\boldsymbol{\mathcal{X}}} + \|\mathbf{F}(t,\boldsymbol{\varphi}) - \mathbf{F}(t,\boldsymbol{\psi})\|_{\boldsymbol{\mathcal{X}}} \\
&\le \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1} (\boldsymbol{L}_{\boldsymbol{\Lambda}}^{\boldsymbol{\mathcal{Y}}} \boldsymbol{C}_{\mathbf{F}} + \boldsymbol{L}_{\mathbf{F}}) \|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}}.
\end{aligned}$$

This concludes the proof. □

Finally, we introduce operators relating the function spaces of the continuous and the discrete problem. These relations are illustrated in Figure 4.1.

($\boldsymbol{\mathcal{J}}$) Let $\boldsymbol{\mathcal{J}} : \mathcal{Y} \to \boldsymbol{\mathcal{X}}$ be a bounded linear operator with

$$\|\boldsymbol{\mathcal{J}}\|_{\mathcal{L}(\mathcal{Y},\boldsymbol{\mathcal{X}})} \le \boldsymbol{C}_{\boldsymbol{\mathcal{J}}}. \tag{4.17}$$

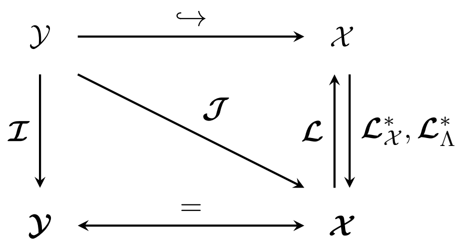

Figure 4.1: Overview of discrete and continuous spaces and operators

($\boldsymbol{\mathcal{I}}$) Let $\boldsymbol{\mathcal{I}} : \mathcal{Y} \to \boldsymbol{\mathcal{Y}}$ be a bounded operator with

$$\|\boldsymbol{\mathcal{I}}\|_{\mathcal{L}(\mathcal{Y},\boldsymbol{\mathcal{Y}})} \leq \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}. \tag{4.18}$$

Note that this condition is stronger than (4.17), as the norm of $\boldsymbol{\mathcal{Y}}$ is in general stronger than the one of $\boldsymbol{\mathcal{X}}$.

($\boldsymbol{\mathcal{L}}$) Let $\boldsymbol{\mathcal{L}} : \boldsymbol{\mathcal{X}} \to \mathcal{X}$ be a bounded linear operator with

$$\|\boldsymbol{\mathcal{L}}\|_{\mathcal{L}(\boldsymbol{\mathcal{X}},\mathcal{X})} \leq \boldsymbol{C}_{\boldsymbol{\mathcal{L}}}. \tag{4.19}$$

As we consider nonconforming space discretizations, we employ $\boldsymbol{\mathcal{L}}$ to map the discrete functions of $\boldsymbol{\mathcal{X}}$ to their continuous counterparts in $\mathcal{X}$. Hence, we call $\boldsymbol{\mathcal{L}}$ the lift operator.

($\boldsymbol{\mathcal{L}}^*_{\mathcal{X}}$) Let $\boldsymbol{\mathcal{L}}^*_{\mathcal{X}} : \mathcal{X} \to \boldsymbol{\mathcal{X}}$ be the adjoint of the lift operator with respect to the standard inner products in $\mathcal{X}$ and $\boldsymbol{\mathcal{X}}$, respectively, i.e.,

$$(\boldsymbol{\mathcal{L}}\boldsymbol{\varphi} \mid \psi)_{\mathcal{X}} = (\boldsymbol{\varphi} \mid \boldsymbol{\mathcal{L}}^*_{\mathcal{X}}\psi)_{\boldsymbol{\mathcal{X}}}\,, \qquad \boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}},\ \psi \in \mathcal{X}. \tag{4.20}$$

($\boldsymbol{\mathcal{L}}^*_{\Lambda}$) For $\xi \in B_{\mathcal{Y}}(R)$ with $\boldsymbol{\mathcal{I}}\xi \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$, let $\boldsymbol{\mathcal{L}}^*_{\Lambda}[\xi] : \mathcal{X} \to \boldsymbol{\mathcal{X}}$ be the adjoint of the lift operator with respect to the weighted inner products $(\cdot \mid \cdot)_{\Lambda(\xi)}$ and $(\cdot \mid \cdot)_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)}$, i.e.,

$$(\boldsymbol{\mathcal{L}}\boldsymbol{\varphi} \mid \psi)_{\Lambda(\xi)} = (\boldsymbol{\varphi} \mid \boldsymbol{\mathcal{L}}^*_{\Lambda}[\xi]\psi)_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)}\,, \qquad \boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}},\ \psi \in \mathcal{X}. \tag{4.21}$$

Before deriving further properties of these operators, we briefly discuss the purpose of these operators with respect to the specific examples. In all examples considered, $\boldsymbol{\mathcal{I}}$ corresponds to an interpolation operator. Moreover, $\boldsymbol{\mathcal{L}}$ is called lift operator, as it is used to lift discrete functions in $\boldsymbol{\mathcal{X}}$ to the continuous space $\mathcal{X}$. Conversely, for the special case of a conforming discretization, the adjoint lift operators are projections from $\mathcal{X}$ to $\boldsymbol{\mathcal{X}}$. Finally, the reference operator $\boldsymbol{\mathcal{J}}$ is used to relate the continuous solution to the discrete framework. For first-order wave-type equations, we choose $\boldsymbol{\mathcal{J}} = \boldsymbol{\mathcal{I}}$. However, for second-order wave-type equations, we have to incorporate the adjoint lift operator $\boldsymbol{\mathcal{L}}^*_{\mathcal{X}}$ in order to prove the expected order of convergence.

As a consequence of (4.19), the adjoint lift operators are also bounded, i.e., we have for arbitrary $\varphi \in \mathcal{X}$ with (4.20)

$$\|\boldsymbol{\mathcal{L}}^*_{\mathcal{X}}\varphi\|_{\boldsymbol{\mathcal{X}}} = \sup_{\|\boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}}=1} (\varphi \mid \boldsymbol{\mathcal{L}}\boldsymbol{\psi})_{\mathcal{X}} \leq \boldsymbol{C}_{\boldsymbol{\mathcal{L}}}\|\varphi\|_{\mathcal{X}}. \tag{4.22}$$

Similarly, we obtain for $\varphi \in \mathcal{X}$ and $\xi \in B_{\mathcal{Y}}(R)$ with $\boldsymbol{\mathcal{I}}\xi \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$ arbitrary using (4.10) as well as (4.21)

$$\|\boldsymbol{\mathcal{L}}^*_\Lambda[\xi]\varphi\|_{\boldsymbol{\mathcal{X}}} \leq \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-\frac{1}{2}} \sup_{\|\psi\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)}=1} (\varphi \mid \boldsymbol{\mathcal{L}}\boldsymbol{\psi})_{\Lambda(\xi)} \leq \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1} \boldsymbol{C}_{\boldsymbol{\Lambda}} \boldsymbol{C}_{\boldsymbol{\mathcal{L}}} \|\varphi\|_{\mathcal{X}}. \tag{4.23}$$

In the next section, we review the basics of semigroup theory for nonautonomous Cauchy problems. We point out that this is essential for the error analysis, as the perturbed evolution equation for the discretization error, which is derived in Section 4.3, belongs to this problem class.

4.2 Abstract evolution equations and semigroups

As a preparation for the error analysis, we first present a brief excursion to semigroups for nonautonomous Cauchy problems based on [Kato, 1970], [Pazy, 1983], [Engel and Nagel, 2000], [Jacob and Zwart, 2012], and the lecture notes [Schnaubelt, 2019], i.e., we consider linear problems of the form

$$\begin{cases} \partial_t z(t) = \mathcal{B}(t) z(t) + g(t), & t \in J_T, \\ z(0) = z_0, \end{cases} \tag{4.24}$$

where, for some Hilbert space $\mathcal{H}$, $z : J_T \to \mathcal{H}$ is the unknown solution, $z_0 \in \mathcal{H}$ is the initial value, $\mathcal{B}$ is a time-dependent linear operator, and $g : J_T \to \mathcal{H}$ is a given right-hand side.

In the following section, we first focus on the case where $\mathcal{B}$ is time invariant. In the second part, we show how the concepts introduced previously can be transferred to problems with time-dependent $\mathcal{B}$. Throughout both sections, let $(\mathcal{H}, (\cdot \mid \cdot)_{\mathcal{H}})$ and $(\mathcal{G}, (\cdot \mid \cdot)_{\mathcal{G}})$ be real-valued Hilbert spaces with corresponding norms $\|\cdot\|_{\mathcal{H}}$ and $\|\cdot\|_{\mathcal{G}}$, respectively. For the sake of presentation, we consider only real-valued Hilbert spaces in this thesis, as this is sufficient for the specific examples. Hence, the results presented in the following chapters are special cases of the corresponding results in the references, where the more general case of complex-valued Hilbert spaces is considered.

4.2.1 Semigroups for Cauchy problems with time-invariant operators

We consider linear Cauchy problems of the form

$$\begin{cases} \partial_t z(t) = \mathcal{B} z(t) + g(t), & t \in J_T, \\ z(0) = z_0, \end{cases} \tag{4.25}$$

where $z : J_T \to \mathcal{H}$ denotes the unknown solution. Furthermore, $z_0 \in \mathcal{H}$ is the initial value, $\mathcal{B} : \mathcal{H} \supset D(\mathcal{B}) \to \mathcal{H}$ is a time-invariant linear operator, and $g : J_T \to \mathcal{H}$ is a given right-hand side.

In order to prove the wellposedness of (4.25), we first introduce the concept of strong continuity.

Definition 4.4. *Let $\{T(t)\}_{t\geq 0} \subset \mathcal{L}(\mathcal{H}, \mathcal{G})$ be a one-parameter family. The map $t \mapsto T(t)$ is called strongly continuous from $\mathcal{H}$ to $\mathcal{G}$ if for all $x \in \mathcal{H}$ the map*

$$T(\cdot)x : [0, \infty) \to \mathcal{G}, \quad t \mapsto T(t)x,$$

is continuous. If $\mathcal{G} = \mathcal{H}$, we simply write "strongly continuous in $\mathcal{H}$".

Based on this definition, we further define strongly continuous semigroups.

Definition 4.5. *A one-parameter family $\{T(t)\}_{t\geq 0} \subset \mathcal{L}(\mathcal{H})$ is called strongly continuous semigroup or simply C_0-semigroup if the following conditions hold.*

(i) $T(0) = \mathrm{Id}$ *and* $T(t+s) = T(t)T(s)$, *for all* $t, s \geq 0$.

(ii) The map $t \mapsto T(t)$ *is strongly continuous in* $\mathcal{H}$.

Before we discuss the connection between semigroups and Cauchy problems, we state a basic property of C_0-semigroups.

Lemma 4.6. *Let $\{T(t)\}_{t\geq 0}$ be a strongly continuous semigroup. Then, there are constants $M \geq 1$ and $\omega \in \mathbb{R}$ such that*

$$\|T(t)\|_{\mathcal{L}(\mathcal{H})} \leq M\mathrm{e}^{\omega t}, \qquad t \geq 0.$$

If the previous lemma holds with $M = 1$, we call $\{T(t)\}_{t\geq 0}$ quasi-contractive. If, in addition, we have $\omega = 0$, we call $\{T(t)\}_{t\geq 0}$ a contraction semigroup. The next definition states that every C_0-semigroup is generated by a unique linear operator.

Definition 4.7. *For a strongly continuous semigroup $\{T(t)\}_{t\geq 0}$ we denote the linear operator $\mathcal{B} : D(\mathcal{B}) \to \mathcal{H}$ defined by*

$$\mathcal{B}x = \lim_{t\to 0+} \frac{T(t)x - x}{t} \tag{4.26}$$

as the infinitesimal generator of $\{T(t)\}_{t\geq 0}$, where the domain $D(\mathcal{B})$ of $\mathcal{B}$ contains all $x \in \mathcal{H}$ for which the limit in (4.26) exists.

Based on the previous definition, the following lemma states that, if $\mathcal{B}$ is the infinitesimal generator of a C_0-semigroup, then the mapping $t \mapsto T(t)x$ is differentiable for all $x \in D(\mathcal{B})$.

Lemma 4.8. *If $\{T(t)\}_{t\geq 0}$ is a strongly continuous semigroup with infinitesimal generator $\mathcal{B}$, then we have the following properties.*

(i) For $x \in D(\mathcal{B})$ *and* $t \geq 0$, *we have* $T(t)x \in D(\mathcal{B})$.

(ii) For all $x \in D(\mathcal{B})$ *and all* $t \geq 0$, *we have*

$$\frac{d}{dt}(T(t)x) = \mathcal{B}T(t)x = T(t)\mathcal{B}x.$$

(iii) $\mathcal{B}$ *is a closed operator with dense domain* $D(\mathcal{B}) \subset \mathcal{H}$.

As a consequence of this lemma, we get that if two strongly continuous semigroups have the same infinitesimal generator, the semigroups coincide. Conversely, the infinitesimal generator of every semigroup is uniquely given by Definition 4.7. Hence, there is a one-to-one relation between C_0-semigroups and their infinitesimal generators.

Finally, we are able to state the wellposedness result for the linear Cauchy problem (4.25).

Theorem 4.9. *Let $\mathcal{B}$ be the infinitesimal generator of a strongly continuous semigroup $\{T(t)\}_{t\geq 0}$ and $g \in C^1(J_T, \mathcal{H})$ or $g \in C(J_T, D(\mathcal{B}))$, where $D(\mathcal{B})$ is equipped with the graph norm $\|\cdot\|_{\mathcal{B}} := \|\cdot\|_{\mathcal{H}} + \|\mathcal{B}\cdot\|_{\mathcal{H}}$. For every initial value $z_0 \in D(\mathcal{B})$, there exists a unique solution z of the linear Cauchy problem (4.25) satisfying*

$$z \in C^1(J_T, \mathcal{H}) \cap C(J_T, D(\mathcal{B})).$$

The solution of (4.25) is given by the variation-of-constants formula

$$z(t) = T(t)z_0 + \int_0^t T(t-s)g(s)\,\mathrm{d}s.$$

Thus, we traced the wellposedness of the linear Cauchy problem (4.25) back to the condition of $\mathcal{B}$ being a generator of a C_0-semigroup. In order to decide whether this is true for a given operator $\mathcal{B}$, we cite the Lumer–Phillips theorem for the Hilbert space setting, cf. [Jacob and Zwart, 2012, Thm. 6.1.7].

Theorem 4.10 (Lumer–Phillips)**.** *Let $\mathcal{B}$ be a linear operator on a Hilbert space $\mathcal{H}$. Then, the following assertions are equivalent.*

(i) $\mathcal{B}$ is densely defined and generates a contraction semigroup.

(ii) $\mathcal{B}$ is dissipative and $\operatorname{range}(\lambda - \mathcal{B}) = \mathcal{H}$ *for some $\lambda > 0$.*

4.2.2 Semigroups for Cauchy problems with time-dependent operators

So far we only considered the wellposedness of problems with a time-invariant operator $\mathcal{B}$. However, to derive an error estimate for the space discretization of quasilinear Cauchy problems, we rely on the wellposedness of problems with a time-dependent operator of the form

$$\begin{cases} \partial_t z(t) = \mathcal{B}(t)z(t) + g(t), & t \in (s, T], \\ \quad z(s) = z_0, \end{cases} \tag{4.27}$$

for an initial time $s \in [0, T)$, where $z : [s, T] \to \mathcal{H}$ denotes the unknown solution. Again, the initial value $z_0 \in D(\mathcal{B}(0))$ and the right-hand side $g : J_T \to \mathcal{H}$ are given. However, in contrast to (4.25), the time-invariant operator $\mathcal{B}$ is replaced by a family of time-dependent operators

$$\{\mathcal{B}(t) : D(\mathcal{B}(t)) \subset \mathcal{H} \to \mathcal{H} \mid t \in J_T\}.$$

Recall that $z \in C^1([s, T], \mathcal{H}) \cap C([s, T], \mathcal{H})$ is a classical solution if $z(t) \in D(\mathcal{B}(t))$ holds for all $t \in [s, T]$, which makes the treatment of these problems more involved.

Nevertheless, we first introduce the concept of evolution families, which extends the notion of continuous semigroups, cf. Definition 4.5.

Definition 4.11. *A two-parameter family $\{U(t, s)\}_{T\geq t\geq s\geq 0} \subset \mathcal{L}(\mathcal{H})$ is called an evolution family if the following conditions hold for all $0 \leq s \leq r \leq t \leq T$:*

(i) $U(s, s) = \mathrm{Id}$, and $U(t, r)U(r, s) = U(t, s)$.

(ii) The mapping $(t, s) \mapsto U(t, s)$ is strongly continuous in $\mathcal{H}$.

We can identify every strongly continuous semigroup $\{T(t)\}_{t\geq 0}$ with the corresponding evolution family given by $U(t,s) = T(t-s)$. However, before we can proceed as in the time-invariant case to derive a bound for the operator norm of an evolution family, we first introduce the notion of a stable family of generators.

Definition 4.12. *Let $[t_\alpha, t_\beta] \subset J_T$. A family $\{\mathcal{B}(t) \mid t \in [t_\alpha, t_\beta]\}$ of infinitesimal generators on $\mathcal{H}$ is called stable if there are constants $M \geq 1$ and $\omega \geq 0$ such that the resolvent set*

$$\rho(\mathcal{B}(t)) := \{\lambda \in \mathbb{C} \mid \lambda \operatorname{Id} - \mathcal{B}(t) : D(\mathcal{B}(t)) \to \mathcal{H} \text{ is bijective}\}$$

satisfies $(\omega, \infty) \subset \rho(\mathcal{B}(t))$ for all $t \in [t_\alpha, t_\beta]$, and

$$\|\mathrm{e}^{s_k \mathcal{B}(t_k)} \mathrm{e}^{s_{k-1} \mathcal{B}(t_{k-1})} \dots \mathrm{e}^{s_1 \mathcal{B}(t_1)}\|_{\mathcal{L}(\mathcal{H})} \leq M \mathrm{e}^{\omega(s_k + s_{k-1} + \dots + s_1)}$$

holds for all $s_j \geq 0$ and $t_\alpha \leq t_1 \leq t_2 \leq \dots \leq t_k \leq t_\beta$. In this case, we write $\mathcal{B}(\cdot) \in \operatorname{stab}(\mathcal{H}, M, \omega)$.

A useful criterion to prove that a family of generators on J_T is stable is shown in [Kato, 1970, Prop. 3.4]. As a direct consequence we obtain the following refined version for subintervals $[t_\alpha, t_\beta] \subset J_T$.

Lemma 4.13. *For $[t_\alpha, t_\beta] \subset J_T$ let $\{\|\cdot\|_t \mid t \in [t_\alpha, t_\beta]\}$ be a family of norms on $\mathcal{H}$, which are all equivalent to $\|\cdot\|_{\mathcal{H}}$ with uniform constants $c_{\mathrm{n}}, C_{\mathrm{n}} > 0$, i.e.,*

$$c_{\mathrm{n}} \|x\|_t \leq \|x\|_{\mathcal{H}} \leq C_{\mathrm{n}} \|x\|_t, \qquad t \in [t_\alpha, t_\beta],\ x \in \mathcal{H},$$

and depend continuously on time in the sense that there is a constant $C' \geq 0$ such that

$$\|x\|_t \leq \mathrm{e}^{C'|t-s|} \|x\|_s, \qquad t, s \in [t_\alpha, t_\beta],\ x \in \mathcal{H}.$$

We denote by $\mathcal{H}_t$ the space $\mathcal{H}$ endowed with $\|\cdot\|_t$. Let further $\{\mathcal{B}(t) \mid t \in [t_\alpha, t_\beta]\}$ be a family of infinitesimal generators of a quasi-contractive semigroup on $\mathcal{H}_t$ such that there exists a constant $\omega \geq 0$ with

$$\|\mathrm{e}^{s\mathcal{B}(t)}\|_{\mathcal{L}(\mathcal{H}_t)} \leq \mathrm{e}^{\omega s}, \qquad s, t \geq 0.$$

Then, we have $\mathcal{B}(\cdot) \in \operatorname{stab}(\mathcal{H}, \frac{C_{\mathrm{n}}}{c_{\mathrm{n}}} \mathrm{e}^{2C'(t_\beta - t_\alpha)}, \omega)$.

With these preliminaries at hand, we can establish the connection between strongly continuous semigroups and the homogeneous, linear, nonautonomous Cauchy problem

$$\begin{cases} \partial_t z(t) = \mathcal{B}(t) z(t), & t \in [s, T], \\ \quad z(s) = z_0. \end{cases} \tag{4.28}$$

To do so, we apply the following refined version of [Kato, 1970, Thm. 4.1, Prop. 6.1] for subintervals $[t_\alpha, t_\beta] \subset J_T$. Under suitable assumptions, this states that the family $\{\mathcal{B}(t) \mid t \in J_T\}$ generates a unique evolution family.

Theorem 4.14. *For $[t_\alpha, t_\beta] \subset J_T$ let $\mathcal{G} \subset \mathcal{H}$ be densely and continuously embedded in $\mathcal{H}$ and $\{\mathcal{B}(t) \mid t \in [t_\alpha, t_\beta]\}$ be a family of generators on $\mathcal{H}$ with the following properties.*

(H_1) There are constants $M \geq 1$ and $\omega \geq 0$ such that $\mathcal{B}(\cdot) \in \operatorname{stab}(\mathcal{H}, M, \omega)$.

(H_2) There exists a family $\{S(t) : \mathcal{G} \to \mathcal{H} \mid t \in [t_\alpha, t_\beta]\}$ of isomorphisms such that $S(t)$ is strongly continuously differentiable from $\mathcal{G}$ to $\mathcal{H}$ for every $t \in [t_\alpha, t_\beta]$. Furthermore, there exists a family $\{\widetilde{\mathcal{B}}(t) \mid t \in [t_\alpha, t_\beta]\}$ such that $\widetilde{\mathcal{B}}(t) \in \mathcal{L}(\mathcal{H})$ is strongly continuous in $\mathcal{H}$ for every $t \in [t_\alpha, t_\beta]$, and

$$S(t)\mathcal{B}(t)S(t)^{-1} = \mathcal{B}(t) + \widetilde{\mathcal{B}}(t), \qquad t \in [t_\alpha, t_\beta].$$

(H_3) $\mathcal{G} \subset D(\mathcal{B}(t))$ and $\mathcal{B}(t) \in \mathcal{L}(\mathcal{G}, \mathcal{H})$ hold for each $t \in [t_\alpha, t_\beta]$. Furthermore, the mapping $t \mapsto \mathcal{B}(t), [t_\alpha, t_\beta] \to \mathcal{L}(\mathcal{G}, \mathcal{H})$ is continuous.

Then, there exists a unique evolution family $\{U(t,s)\}_{t_\beta \geq t \geq s \geq t_\alpha} \subset \mathcal{L}(\mathcal{H})$ satisfying

(E_1) $\|U(t,s)\|_{\mathcal{L}(\mathcal{H})} \leq M \mathrm{e}^{\omega(t-s)}$, for all $t_\alpha \leq s \leq t \leq t_\beta$.

(E_2) For $t_\alpha \leq s \leq t \leq t_\beta$, $U(t,s)$ is strongly continuous in $\mathcal{G}$ both in s and t.

(E_3) For every $x \in \mathcal{G}$ and $s \in [t_\alpha, t_\beta)$, we have that $U(\cdot, s)x : [s, t_\beta] \to \mathcal{H}$ is strongly continuously differentiable in $\mathcal{H}$ with

$$\frac{\partial}{\partial t} U(t,s)x = \mathcal{B}(t)U(t,s)x.$$

Hence, we get the following wellposedness result, cf. [Pazy, 1983, Thm. 4.3].

Theorem 4.15. *If $\mathcal{H}$, $\mathcal{G}$ and $\{\mathcal{B}(t) \mid t \in J_T\}$ satisfy the assumptions of Theorem 4.14, then (4.28) has for every $s \in [0, T)$ a unique solution z satisfying*

$$z \in C^1([s,T], \mathcal{H}) \cap C([s,T], \mathcal{G}).$$

Finally, we link the concept of evolution families to the wellposedness of the inhomogeneous, linear Cauchy problem with a time-dependent operator (4.27), cf. [Kato, 1970, Thm. 7.1].

Theorem 4.16. *Assume $g \in C(J_T, \mathcal{G})$ and that $\mathcal{G}$, $\mathcal{H}$ and $\{\mathcal{B}(t) \mid t \in J_T\}$ satisfy the assumptions of Theorem 4.14, i.e., there exists a unique evolution family $\{U(t,s)\}_{T \geq t \geq s \geq 0} \subset \mathcal{L}(\mathcal{H})$. For every $s \in [0, T)$ and any initial value $z_0 \in \mathcal{G}$, there exists a unique solution z of (4.27) with*

$$z \in C^1([s,T], \mathcal{H}) \cap C([s,T], \mathcal{G}),$$

which is given by the variation-of-constants formula

$$z(t) = U(t,s)z_0 + \int_s^t U(t,r)g(r)\,\mathrm{d}r, \qquad t \in [s,T].$$

Making use of the isomorphisms from (H_2), the authors of [Hochbruck and Pažur, 2017], [Hochbruck et al., 2018], and [Kovács and Lubich, 2018] provide wellposedness and an error analysis for the application of Runge–Kutta methods to quasilinear wave-type problems. We briefly review these results in Section 6.1.

Hence, it would be intuitive to follow the same approach to analyze the space discretization as well as the full discretization. However, this is problematic, as discussed in the following remark.

Remark 4.17. *Up to our knowledge, it is not clear whether a family of isomorphisms as used in (H_2) for general discrete spaces $\mathcal{H} = \boldsymbol{\mathcal{X}}$ and $\mathcal{G} = \boldsymbol{\mathcal{Y}}$ exists. Nevertheless, we can still apply the results for the choice $\mathcal{H} = \boldsymbol{\mathcal{X}} = \mathcal{G}$, as these are finite-dimensional spaces, and hence all operators are bounded.*

The important point is that, although the operator norm in (H_3) then depends on the discretization parameter, this dependency does not affect the result. In fact, keeping track of the norm of $\mathcal{B}$ in the proofs of [Kato, 1970, Thm. 4.1, Prop. 6.1], we see that the continuity in time is simply necessary to justify the approximation of $\{\mathcal{B}(t) \mid t \in J_T\}$ by a sequence of piecewise constant operators, where the magnitude of the operator norm affects the rate of convergence of this approximation, but not the limit.

However, the choice $\mathcal{H} = \boldsymbol{\mathcal{X}} = \mathcal{G}$ is only feasible to derive an error estimate for the quasilinear wave-type problem (4.12), since we require distinct spaces $\mathcal{H} = \boldsymbol{\mathcal{X}}$ and $\mathcal{G} = \boldsymbol{\mathcal{Y}}$ for the wellposedness analysis. Moreover, it is essential that the operator norm of the isomorphisms in (H_2) is independent of the spatial discretization parameter. Hence, it is not possible to apply the theory as in the analysis of time discretizations.

As a final comment, we point out that for $\mathcal{B}(t)$ being a bounded operator independent of t, the wellposedness of (4.27) is also shown in [Pazy, 1983, Chap. 5.1, Thm. 5.1]. Due to the additional assumption, the proof of this result is much simpler compared to the proofs of [Kato, 1970, Thm. 4.1, Prop. 6.1]. However, note that this result is not sufficient for the analysis of the space discretization below, as it does not include the bound from (E_1) with uniform constants M, ω, which is essential to derive an error estimate in the next section.

4.3 Analysis of the abstract space discretization

In this section, we show the wellposedness of the discrete quasilinear Cauchy problem (4.12) and derive a rigorous error estimate. Usually, the first step in the error analysis would be to prove wellposedness. Then, once the unique existence of both the continuous and the discrete solution is known, the next step would be to derive an error estimate. However, in our case it is not possible to follow the standard approach, as the proofs of wellposedness and convergence are intertwined here. To be more precise, the main difficulty is that it is not sufficient to prove existence of a discrete solution, which is bounded in $\boldsymbol{\mathcal{X}}$, although this is the natural space for the problem. Instead, we have to provide bounds in the $\boldsymbol{\mathcal{Y}}$-norm in order to prevent degeneracy of the problem, i.e., the properties of the discrete operators stated in Assumption 4.1 are only valid under these bounds.

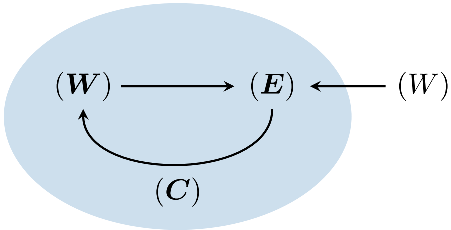

Figure 4.2: Roadmap for the analysis of the abstract space discretization.

Roadmap to prove wellposedness and error bounds

The general roadmap consists of the following steps.

(W) By Assumption 3.2, the continuous quasilinear Cauchy problem (3.3) is wellposed. In particular, for $T < t^*(y_0)$ this ensures the existence of a unique solution y of (3.3) on J_T, which satisfies $\|y\|_{\mathcal{Y}} < R$ uniformly on J_T, where R is the radius used in Assumption 3.1.

($\boldsymbol{W}$) In Lemma 4.18 we prove that there exists a unique solution $\boldsymbol{y}$ of the discrete quasilinear Cauchy problem (4.12) on a time interval J_{T_h}. In particular, we show that this solution satisfies $\|\boldsymbol{y}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}$ uniformly on J_{T_h}, where the radius $\boldsymbol{R}$ is given by Assumption 4.1. To prove this result, we employ that $\boldsymbol{\mathcal{Y}}$ is a finite-dimensional space and apply the Picard–Lindelöf theorem. However, as the Lipschitz constant of the right-hand side of the equation depends on the discretization parameter h, we get $T_h \to 0$ for $h \to 0$.

($\boldsymbol{E}$) In Theorem 4.20 we estimate the error $y - \boldsymbol{\mathcal{L}}\boldsymbol{y}$ in the weaker $\mathcal{X}$-norm on the time interval $\widetilde{J} = [0, \min\{T, T_h\}]$, using the semigroup theory presented in the previous section.

($\boldsymbol{C}$) With Assumption 4.22 we conclude that $\|\boldsymbol{y} - \boldsymbol{\mathcal{I}}y\|_{\boldsymbol{\mathcal{Y}}} \to 0$ uniformly on $\widetilde{J}$, for $h \to 0$. This allows us to prove $T_h \geq T$ for h sufficiently small (Theorem 4.25).

This approach is also illustrated in Figure 4.2, where the analysis of the discrete quasilinear Cauchy problem is indicated by the blue ellipse.

Note that, since we prove these results uniformly for $T \in (0, t^*(y_0))$, we get the same estimates also for the maximal times of existence $t^*(y_0)$ and $t^*(\boldsymbol{y}_0)$.

In the following lemma, we start by proving the wellposedness of the discrete quasilinear Cauchy problem (4.12). Since $\boldsymbol{\mathcal{X}}$ is a finite-dimensional space, there is a constant $\boldsymbol{C}_{\mathbf{A}}(h) > 0$ such that

$$\|\mathbf{A}\|_{\mathcal{L}(\boldsymbol{\mathcal{X}})} \leq \boldsymbol{C}_{\mathbf{A}}(h) \tag{4.29}$$

holds, i.e., $\mathbf{A}$ is bounded. Although the constant in this bound may deteriorate for $h \to 0$, this nevertheless allows us to apply the theorem of Picard–Lindelöf. Finally, note that we use the local version of Picard–Lindelöf here, as we have to bound the solution in $\boldsymbol{\mathcal{Y}}$.

Lemma 4.18. *Let $\boldsymbol{R} > 0$ be the radius from Assumption 4.1 and $\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} > 0$ be arbitrary. If*

$$\|\boldsymbol{y}_0\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}, \qquad \|\boldsymbol{\mathcal{A}}(\boldsymbol{y}_0)\boldsymbol{y}_0\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}^{\boldsymbol{\mathcal{A}}},$$

then there exists a maximal time of existence $t_h^*(\boldsymbol{y}_0) > 0$ such that for all $T_h < t_h^*(\boldsymbol{y}_0)$, (4.12) has a unique solution $\boldsymbol{y}$ which satisfies

$$\boldsymbol{y} \in C^1(J_{T_h}, \boldsymbol{\mathcal{X}}) \cap C(J_{T_h}, \boldsymbol{\mathcal{Y}}) \tag{4.30}$$

with

$$\|\boldsymbol{y}(t)\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}, \qquad \|\boldsymbol{\mathcal{A}}(\boldsymbol{y}(t))\boldsymbol{y}(t)\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}^{\boldsymbol{\mathcal{A}}}, \qquad t \in J_{T_h}. \tag{4.31}$$

Proof. We show that the right-hand side of (4.12) is Lipschitz continuous with respect to the second argument. Let $\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2 \in \boldsymbol{\mathcal{Y}}$ with

$$\|\boldsymbol{\varphi}_i\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}, \qquad \|\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}_i)\boldsymbol{\varphi}_i\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}^{\boldsymbol{\mathcal{A}}}, \qquad i = 1, 2.$$

The triangle inequality yields

$$\|\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}_1)\boldsymbol{\varphi}_1 - \boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}_2)\boldsymbol{\varphi}_2\|_{\boldsymbol{\mathcal{X}}} \le \|\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}_1)(\boldsymbol{\varphi}_1 - \boldsymbol{\varphi}_2)\|_{\boldsymbol{\mathcal{X}}} + \|(\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}_1) - \boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}_2))\boldsymbol{\varphi}_2\|_{\boldsymbol{\mathcal{X}}}.$$

With (4.13) as well as the bounds (4.16) and (4.29) for $\boldsymbol{\Lambda}(\boldsymbol{\varphi}_1)^{-1}$ and $\mathbf{A}$, respectively, we further get

$$\|\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}_1)(\boldsymbol{\varphi}_1 - \boldsymbol{\varphi}_2)\|_{\boldsymbol{\mathcal{X}}} \le \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1} \boldsymbol{C}_{\mathbf{A}}(h)\|\boldsymbol{\varphi}_1 - \boldsymbol{\varphi}_2\|_{\boldsymbol{\mathcal{X}}}.$$

The Lipschitz continuity (4.14) together with $\|\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}_2)\boldsymbol{\varphi}_2\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}^{\boldsymbol{\mathcal{A}}}$ yields

$$\begin{aligned}\|(\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}_1) - \boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}_2))\boldsymbol{\varphi}_2\|_{\boldsymbol{\mathcal{X}}} &\le \boldsymbol{L}_{\boldsymbol{\mathcal{A}}}\|\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}_2)\boldsymbol{\varphi}_2\|_{\boldsymbol{\mathcal{Y}}}\|\boldsymbol{\varphi}_1 - \boldsymbol{\varphi}_2\|_{\boldsymbol{\mathcal{X}}} \\ &\le \boldsymbol{L}_{\boldsymbol{\mathcal{A}}}\boldsymbol{R}^{\boldsymbol{\mathcal{A}}}\|\boldsymbol{\varphi}_1 - \boldsymbol{\varphi}_2\|_{\boldsymbol{\mathcal{X}}}.\end{aligned}$$

For $t \in J_T$, we further get from (4.15)

$$\|\boldsymbol{\mathcal{F}}(t, \boldsymbol{\varphi}_1) - \boldsymbol{\mathcal{F}}(t, \boldsymbol{\varphi}_2)\|_{\boldsymbol{\mathcal{X}}} \le \boldsymbol{L}_{\boldsymbol{\mathcal{F}}}\|\boldsymbol{\varphi}_1 - \boldsymbol{\varphi}_2\|_{\boldsymbol{\mathcal{X}}}.$$

Collecting these results and using the triangle inequality together with the inverse estimates (4.1), we finally have shown

$$\|\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}_1)\boldsymbol{\varphi}_1 + \boldsymbol{\mathcal{F}}(t, \boldsymbol{\varphi}_1) - (\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}_2)\boldsymbol{\varphi}_2 + \boldsymbol{\mathcal{F}}(t, \boldsymbol{\varphi}_2))\|_{\boldsymbol{\mathcal{Y}}} \le C\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)(1 + \boldsymbol{C}_{\mathbf{A}}(h))\boldsymbol{C}_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}}(h)\|\boldsymbol{\varphi}_1 - \boldsymbol{\varphi}_2\|_{\boldsymbol{\mathcal{Y}}}$$

with a constant C depending both on $\boldsymbol{R}^{\boldsymbol{\mathcal{A}}}$ and the constants from Assumption 4.1, but not on h. Hence, the Picard–Lindelöf theorem yields the result. □

Since we use estimates which depend on the discretization parameter, it is not surprising that we are only able to guarantee the existence of the solution up to some time $T_h < t_h^*(\boldsymbol{y}_0)$, which again depends on the discretization. However, we show at the end of this section that, under suitable assumptions on the discretization, we get $t_h^*(\boldsymbol{y}_0) \ge t^*(y_0)$, i.e., the discrete approximation obtained by (4.12) exists at least as long as the solution of (3.3).

Motivated by the unified error analysis proposed in [Hipp et al., 2019] for linear wave-type problems and [Hochbruck and Leibold, 2019] for semi-linear wave-type problems, the main result in this section states an estimate for the error between the exact solution and the lifted solution of the semidiscrete problem. To do so, we assume in the following that the radii R and $\boldsymbol{R}$ from Assumptions 3.1 and 4.1, respectively, are chosen such that

$$\boldsymbol{C}_{\boldsymbol{\mathcal{I}}} R < \boldsymbol{R}. \tag{4.32}$$

We employ the following definition of the remainder terms.

$$\mathcal{R}_{\Lambda}(\xi) \coloneqq \boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_{\mathcal{X}}^{*}\Lambda(\xi), \qquad \xi \in B_{\mathcal{Y}}(R), \tag{4.33}$$

$$\mathcal{R}_{\mathrm{A}} \coloneqq \mathbf{A}\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_{\mathcal{X}}^{*}\mathrm{A}, \tag{4.34}$$

$$\mathcal{R}_{\mathrm{F}}(t,\xi) \coloneqq \mathbf{F}(t,\boldsymbol{\mathcal{I}}\xi) - \boldsymbol{\mathcal{L}}_{\mathcal{X}}^{*}\mathrm{F}(t,\xi), \qquad t \in J_T,\ \xi \in B_{\mathcal{Y}}(R). \tag{4.35}$$

Note that $\mathcal{R}_{\Lambda}(\xi)$ and $\mathcal{R}_{\mathrm{F}}(t,\xi)$ are well defined if R and $\boldsymbol{R}$ in Assumptions 3.1 and 4.1, respectively, are chosen such that (4.32) holds, due to the boundedness (4.18) of $\boldsymbol{\mathcal{I}}$. As these remainders also occur in the full discretization, we prove here a preliminary lemma for the remainders, which is also needed in the proof of the main result.

Lemma 4.19. *For $t \in J_T$, $\zeta \in \mathcal{Y}$, and $\xi \in B_{\mathcal{Y}}(R)$, we have*

$$\|(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_{\Lambda}^{*}[\xi])\zeta\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)} \leq \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-\frac{1}{2}} \|\mathcal{R}_{\Lambda}(\xi)\zeta\|_{\boldsymbol{\mathcal{X}}}, \tag{4.36}$$

$$\|(\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}\xi)\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_{\Lambda}^{*}[\xi]\mathcal{A}(\xi))\zeta\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)} \leq \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-\frac{1}{2}} \|\mathcal{R}_{\mathrm{A}}\zeta\|_{\boldsymbol{\mathcal{X}}}, \tag{4.37}$$

$$\|\boldsymbol{\mathcal{F}}(t,\boldsymbol{\mathcal{I}}\xi) - \boldsymbol{\mathcal{L}}_{\Lambda}^{*}[\xi]\mathcal{F}(t,\xi)\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)} \leq \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-\frac{1}{2}} \|\mathcal{R}_{\mathrm{F}}(t,\xi)\|_{\boldsymbol{\mathcal{X}}}. \tag{4.38}$$

Proof. Let $t \in J_T$, $\zeta \in \mathcal{Y}$, $\xi \in B_{\mathcal{Y}}(R)$ and $\boldsymbol{\zeta} \in \boldsymbol{\mathcal{X}}$ arbitrary. For (4.36), we use the definitions of the adjoint lift operators (4.20) and (4.21) together with the definitions of the state-dependent inner products (3.5) and (4.8). This yields

$$\begin{aligned}
((\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_{\Lambda}^{*}[\xi])\zeta \mid \boldsymbol{\zeta})_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)} &= (\boldsymbol{\mathcal{J}}\zeta \mid \boldsymbol{\zeta})_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)} - (\zeta \mid \boldsymbol{\mathcal{L}}\boldsymbol{\zeta})_{\Lambda(\xi)} \\
&= (\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)\boldsymbol{\mathcal{J}}\zeta \mid \boldsymbol{\zeta})_{\boldsymbol{\mathcal{X}}} - (\Lambda(\xi)\zeta \mid \boldsymbol{\mathcal{L}}\boldsymbol{\zeta})_{\mathcal{X}} \\
&= (\mathcal{R}_{\Lambda}(\xi)\zeta \mid \boldsymbol{\zeta})_{\boldsymbol{\mathcal{X}}}\,.
\end{aligned}$$

To prove the corresponding bound (4.37), we get with (3.4) and (4.13)

$$\begin{aligned}
((\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}\xi)\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_{\Lambda}^{*}[\xi]\mathcal{A}(\xi))\zeta \mid \boldsymbol{\zeta})_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)} &= (\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}\xi)\boldsymbol{\mathcal{J}}\zeta \mid \boldsymbol{\zeta})_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)} - (\mathcal{A}(\xi)\zeta \mid \boldsymbol{\mathcal{L}}\boldsymbol{\zeta})_{\Lambda(\xi)} \\
&= (\mathbf{A}\boldsymbol{\mathcal{J}}\zeta \mid \boldsymbol{\zeta})_{\boldsymbol{\mathcal{X}}} - (\boldsymbol{\mathcal{L}}_{\mathcal{X}}^{*}\mathrm{A}\zeta \mid \boldsymbol{\zeta})_{\boldsymbol{\mathcal{X}}} \\
&= (\mathcal{R}_{\mathrm{A}}\zeta \mid \boldsymbol{\zeta})_{\boldsymbol{\mathcal{X}}}\,.
\end{aligned}$$

Using the same arguments, we derive for (4.38)

$$\begin{aligned}
(\boldsymbol{\mathcal{F}}(t,\boldsymbol{\mathcal{I}}\xi) - \boldsymbol{\mathcal{L}}_{\Lambda}^{*}[\xi]\mathcal{F}(t,\xi) \mid \boldsymbol{\zeta})_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)} &= (\boldsymbol{\mathcal{F}}(t,\boldsymbol{\mathcal{I}}\xi) \mid \boldsymbol{\zeta})_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)} - (\mathcal{F}(t,\xi) \mid \boldsymbol{\mathcal{L}}\boldsymbol{\zeta})_{\Lambda(\xi)} \\
&= (\mathbf{F}(t,\boldsymbol{\mathcal{I}}\xi) \mid \boldsymbol{\zeta})_{\boldsymbol{\mathcal{X}}} - (\boldsymbol{\mathcal{L}}_{\mathcal{X}}^{*}\mathrm{F}(t,\xi) \mid \boldsymbol{\zeta})_{\boldsymbol{\mathcal{X}}} \\
&= (\mathcal{R}_{\mathrm{F}}(t,\xi) \mid \boldsymbol{\zeta})_{\boldsymbol{\mathcal{X}}}\,.
\end{aligned}$$

Finally, we use

$$\|\boldsymbol{\varphi}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)} = \sup_{\|\boldsymbol{\zeta}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)}=1} (\boldsymbol{\varphi} \mid \boldsymbol{\zeta})_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)}\,, \qquad \boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}},$$

together with the Cauchy–Schwarz inequality and the norm equivalence (4.10) to deduce the results. □

With these results at hand, we now state the main result of this section.

Theorem 4.20. *Let Assumptions 3.1 and 4.1 be true with radii $R, \boldsymbol{R} > 0$, which satisfy (4.32). Let y be the solution of (3.3), which satisfies*

$$y \in C^1([0,t^*(y_0)),\mathcal{Y}) \cap C([0,t^*(y_0)), B_{\mathcal{Y}}(R)).$$

Further, let the assumptions of Lemma 4.18 be satisfied, i.e., there exists a unique solution $\boldsymbol{y}$ of (4.12) with maximal time of existence $t_h^*(\boldsymbol{y}_0)$ satisfying (4.31). If for $T < \min\{t^*(y_0), t_h^*(\boldsymbol{y}_0)\}$ there is a radius $R^{\partial_t} > 0$ such that

$$\|\partial_t y(t)\|_{\mathcal{Y}} < R^{\partial_t}, \qquad t \in J_T,$$

holds, then for $t \in J_T$ the error satisfies

$$\begin{aligned}\|y(t) - \boldsymbol{\mathcal{L}}\boldsymbol{y}(t)\|_{\mathcal{X}} \le \|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{J}})y(t)\|_{\mathcal{X}} + C(1+t)\mathrm{e}^{Ct}\Big(&\|\boldsymbol{\mathcal{J}}y_0 - \boldsymbol{y}_0\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t]}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})y\|_{\boldsymbol{\mathcal{X}}} \\ &+ \sup_{[0,t]}\|\mathcal{R}_\Lambda(y)\partial_t y\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t]}\|\mathcal{R}_A y\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t]}\|\mathcal{R}_F(\cdot,y)\|_{\boldsymbol{\mathcal{X}}}\Big),\end{aligned} \tag{4.39}$$

with a constant $C > 0$, which depends on R^{∂_t}, $t^*(y_0)$, and the radius $\boldsymbol{R}^{\boldsymbol{\mathcal{A}}}$ from (4.31), but is independent of h and T.

Proof. Let $t < \min\{t^*(y_0), t_h^*(\boldsymbol{y}_0)\}$. We first split the error into

$$y(t) - \boldsymbol{\mathcal{L}}\boldsymbol{y}(t) = (\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{J}})y(t) + \boldsymbol{\mathcal{L}}(\boldsymbol{\mathcal{J}}y(t) - \boldsymbol{y}(t)).$$

With the discretization error

$$\boldsymbol{e}(t) = \boldsymbol{\mathcal{J}}y(t) - \boldsymbol{y}(t)$$

and the boundedness (4.19) of the lift operator, we obtain

$$\|y(t) - \boldsymbol{\mathcal{L}}\boldsymbol{y}(t)\|_{\mathcal{X}} \le \|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{J}})y(t)\|_{\mathcal{X}} + \boldsymbol{C}_{\boldsymbol{\mathcal{L}}}\|\boldsymbol{e}(t)\|_{\boldsymbol{\mathcal{X}}}.$$

The first term appears in the right-hand side of (4.39). Hence, it is sufficient to focus on second term. By (4.12), $\boldsymbol{e}$ satisfies the evolution equation

$$\begin{aligned}\partial_t\boldsymbol{e}(t) &= \boldsymbol{\mathcal{J}}\partial_t y(t) - \partial_t\boldsymbol{y}(t) \\ &= \boldsymbol{\mathcal{J}}\partial_t y(t) - \boldsymbol{\mathcal{A}}(\boldsymbol{y}(t))\boldsymbol{y}(t) - \boldsymbol{\mathcal{F}}(t,\boldsymbol{y}(t)) \\ &= -\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y(t))\boldsymbol{\mathcal{J}}y(t) + \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y(t))(\boldsymbol{\mathcal{J}}y(t) - \boldsymbol{y}(t)) + (\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y(t)) - \boldsymbol{\mathcal{A}}(\boldsymbol{y}(t)))\boldsymbol{y}(t) \\ &\quad + \boldsymbol{\mathcal{J}}\partial_t y(t) - \boldsymbol{\mathcal{F}}(t,\boldsymbol{y}(t)).\end{aligned}$$

With

$$\boldsymbol{g}(t) = (\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y(t)) - \boldsymbol{\mathcal{A}}(\boldsymbol{y}(t)))\boldsymbol{y}(t) + \boldsymbol{\mathcal{J}}\partial_t y(t) - \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y(t))\boldsymbol{\mathcal{J}}y(t) - \boldsymbol{\mathcal{F}}(t, \boldsymbol{y}(t)), \tag{4.40}$$

we thus obtain

$$\begin{cases} \partial_t \boldsymbol{e}(t) = \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y(t))\boldsymbol{e}(t) + \boldsymbol{g}(t), \quad t \in J_T, \\ \quad \boldsymbol{e}(0) = \boldsymbol{\mathcal{J}}y_0 - \boldsymbol{y}_0, \end{cases} \tag{4.41}$$

for $T < \min\{t^*(y_0), t_h^*(\boldsymbol{y}_0)\}$. This is a nonautonomous Cauchy problem, which fits into the framework (4.24). Thus, by Theorem 4.16 we have to check assumptions (H_1), (H_2), and (H_3) for $[t_\alpha, t_\beta] = J_T$ to show the wellposedness of (4.41). As explained in Remark 4.17, the idea is to set both $\mathcal{H} = \boldsymbol{\mathcal{X}}$ and $\mathcal{G} = \boldsymbol{\mathcal{X}}$.

To prove (H_1), note that the definition (4.13) of $\boldsymbol{\mathcal{A}}$ together with the dissipativity (4.5) of $\mathbf{A}$ imply that $\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y(t))$ is dissipative with respect to the weighted inner product $(\cdot \mid \cdot)_{\mathbf{A}(\boldsymbol{\mathcal{I}}y(t))}$. As $\boldsymbol{\mathcal{X}}$ is a finite-dimensional space, the Lumer–Phillips theorem (Theorem 4.10) yields that $\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y(t))$ is the infinitesimal generator of a contraction semigroup for every $t \in J_T$. Together with the norm equivalence (4.10) and the norm continuity (4.11), we therefore have shown that Lemma 4.13 is applicable. This yields $\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y(\cdot)) \in \operatorname{stab}(\mathcal{H}, M, 0)$ with a constant $M = \boldsymbol{c}_{\mathbf{\Lambda}}^{-\frac{1}{2}} \boldsymbol{C}_{\mathbf{\Lambda}}^{\frac{1}{2}} \mathrm{e}^{2C'T}$ based on the constants appearing in Lemma 4.2. Thus, (H_1) is satisfied.

Next, we see that (H_2) is satisfied for $S \equiv \mathrm{Id}$ and $\widetilde{\mathcal{B}} \equiv 0$.

Finally, we have to ensure the continuity of $t \mapsto \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y(t))$ as a map from J_T to $\mathcal{L}(\boldsymbol{\mathcal{X}})$. However, we emphasize that the constants used to verify (H_2) do not affect the overall result, as explained in Remark 4.17. Hence, having a constant which depends on the discretization parameter is not an issue here.

Let $s_1, s_2 \geq 0$. First, (4.14) yields

$$\|(\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y(s_2)) - \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y(s_1)))\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}} \leq \boldsymbol{L}_{\boldsymbol{\mathcal{A}}} \|\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y(s_2))\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{Y}}} \|\boldsymbol{\mathcal{I}}(y(s_2) - y(s_1))\|_{\boldsymbol{\mathcal{X}}}.$$

The inverse estimates (4.1), the norm equivalence (4.10), and the boundedness (4.29) of $\mathbf{A}$ yield

$$\|\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y(s_2))\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{Y}}} \leq \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h) \boldsymbol{c}_{\mathbf{\Lambda}}^{-1} \boldsymbol{C}_{\mathbf{A}}(h) \boldsymbol{C}_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}}(h) \boldsymbol{R}.$$

Furthermore, the bound (4.18) on $\boldsymbol{\mathcal{I}}$, and the fundamental theorem of calculus yield

$$\|\boldsymbol{\mathcal{I}}(y(s_2) - y(s_1))\|_{\boldsymbol{\mathcal{X}}} \leq \boldsymbol{C}_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}}(h) \boldsymbol{C}_{\boldsymbol{\mathcal{I}}} R^{\partial_t} |s_2 - s_1|.$$

Collecting the results, we finally have shown

$$\|(\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y(s_2)) - \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y(s_1)))\boldsymbol{\xi}\|_{\mathcal{L}(\boldsymbol{\mathcal{X}})} \leq C(h)|s_2 - s_1|$$

with some constant $C(h) > 0$ depending on the discretization parameter h. This proves (H_2).

Hence, Theorem 4.16 yields the existence of an evolution family $\left(\boldsymbol{U}(t,s)\right)_{T \geq t \geq s \geq 0}$ such that the discrete error is given by

$$\boldsymbol{e}(t) = \boldsymbol{U}(t,0)e(0) + \int_0^t \boldsymbol{U}(t,s)\boldsymbol{g}(s)\,\mathrm{d}s.$$

Furthermore, Lemma 4.13 and Theorem 4.14 for $[t_\alpha, t_\beta] = [s,t]$ yield the estimate

$$\|\boldsymbol{U}(t,s)\|_{\mathcal{L}(\boldsymbol{\mathcal{X}})} \le \left(\frac{\boldsymbol{C_\Lambda}}{\boldsymbol{c_\Lambda}}\right)^{\frac{1}{2}} \mathrm{e}^{2C'(t-s)}, \qquad 0 \le s \le t.$$

This implies

$$\begin{aligned}\|\boldsymbol{e}(t)\|_{\boldsymbol{\mathcal{X}}} &\le \|\boldsymbol{U}(t,0)\boldsymbol{e}(0)\|_{\boldsymbol{\mathcal{X}}} + \int_0^t \|\boldsymbol{U}(t,s)\boldsymbol{g}(s)\|_{\boldsymbol{\mathcal{X}}}\,\mathrm{d}s \\ &\le \left(\frac{\boldsymbol{C_\Lambda}}{\boldsymbol{c_\Lambda}}\right)^{\frac{1}{2}} \mathrm{e}^{2C't}\|\boldsymbol{\mathcal{J}}y_0 - \boldsymbol{y}_0\|_{\boldsymbol{\mathcal{X}}} + \left(\frac{\boldsymbol{C_\Lambda}}{\boldsymbol{c_\Lambda}}\right)^{\frac{1}{2}} \int_0^t \mathrm{e}^{2C'(t-s)}\|\boldsymbol{g}(s)\|_{\boldsymbol{\mathcal{X}}}\,\mathrm{d}s.\end{aligned} \tag{4.42}$$

We now prove a bound for $\|\boldsymbol{g}(s)\|_{\boldsymbol{\mathcal{X}}}$ with $s \in [0,t]$, where we omit the dependency on time whenever possible for the ease of presentation. First, addition of the adjoint lift operator $\boldsymbol{\mathcal{L}}_\Lambda^*[y]$ applied to (3.3) to (4.40) and reordering the terms yields

$$\begin{aligned}\boldsymbol{g} &= (\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y) - \boldsymbol{\mathcal{A}}(\boldsymbol{y}))\boldsymbol{y} + \boldsymbol{\mathcal{J}}\partial_t y - \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y)\boldsymbol{\mathcal{J}}y - \boldsymbol{\mathcal{F}}(\boldsymbol{y}) - \boldsymbol{\mathcal{L}}_\Lambda^*[y](\partial_t y - \mathcal{A}(y)y - \mathcal{F}(y)) \\ &= (\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y) - \boldsymbol{\mathcal{A}}(\boldsymbol{y}))\boldsymbol{y} - \boldsymbol{\mathcal{F}}(\boldsymbol{y}) \\ &\quad + (\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_\Lambda^*[y])\partial_t y + (\boldsymbol{\mathcal{L}}_\Lambda^*[y]\mathcal{A}(y) - \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y)\boldsymbol{\mathcal{J}})y + \boldsymbol{\mathcal{L}}_\Lambda^*[y]\mathcal{F}(y) \\ &= (\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y) - \boldsymbol{\mathcal{A}}(\boldsymbol{y}))\boldsymbol{y} + \boldsymbol{\mathcal{F}}(\boldsymbol{\mathcal{I}}y) - \boldsymbol{\mathcal{F}}(\boldsymbol{y}) \\ &\quad + (\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_\Lambda^*[y])\partial_t y + (\boldsymbol{\mathcal{L}}_\Lambda^*[y]\mathcal{A}(y) - \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y)\boldsymbol{\mathcal{J}})y + \boldsymbol{\mathcal{L}}_\Lambda^*[y]\mathcal{F}(y) - \boldsymbol{\mathcal{F}}(\boldsymbol{\mathcal{I}}y).\end{aligned} \tag{4.43}$$

We bound the terms separately. First, the Lipschitz continuity (4.14) yields

$$\begin{aligned}\|(\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}y) - \boldsymbol{\mathcal{A}}(\boldsymbol{y}))\boldsymbol{y}\|_{\boldsymbol{\mathcal{X}}} &\le \boldsymbol{L_{\mathcal{A}}}\|\boldsymbol{\mathcal{A}}(\boldsymbol{y})\boldsymbol{y}\|_{\boldsymbol{\mathcal{Y}}}\|\boldsymbol{\mathcal{I}}y - \boldsymbol{y}\|_{\boldsymbol{\mathcal{X}}} \\ &\le \boldsymbol{L_{\mathcal{A}}}\boldsymbol{R}^{\boldsymbol{\mathcal{A}}}(\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})y\|_{\boldsymbol{\mathcal{X}}} + \|\mathbf{e}\|_{\boldsymbol{\mathcal{X}}}).\end{aligned}$$

Similarly, for the second difference with the Lipschitz continuity (4.15), we obtain

$$\begin{aligned}\|\boldsymbol{\mathcal{F}}(\boldsymbol{\mathcal{I}}y) - \boldsymbol{\mathcal{F}}(\boldsymbol{y})\|_{\boldsymbol{\mathcal{X}}} &\le \boldsymbol{L_{\mathcal{F}}}\|\boldsymbol{y} - \boldsymbol{\mathcal{I}}y\|_{\boldsymbol{\mathcal{X}}} \\ &\le \boldsymbol{L_{\mathcal{F}}}(\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})y\|_{\boldsymbol{\mathcal{X}}} + \|\mathbf{e}\|_{\boldsymbol{\mathcal{X}}}).\end{aligned}$$

Finally, the last three terms can be bounded by Lemma 4.19 due to the norm equivalence (4.10). Hence, Lemma 4.19 yields

$$\|\boldsymbol{g}(s)\|_{\boldsymbol{\mathcal{X}}} \le C(\|\boldsymbol{e}(s)\|_{\boldsymbol{\mathcal{X}}} + \|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})y(s)\|_{\boldsymbol{\mathcal{X}}} + \|\mathcal{R}_\Lambda(y(s))\partial_t y(s)\|_{\boldsymbol{\mathcal{X}}} + \|\mathcal{R}_\mathrm{A} y(s)\|_{\boldsymbol{\mathcal{X}}} + \|\mathcal{R}_\mathrm{F}(s,y)\|_{\boldsymbol{\mathcal{X}}})$$

with a constant $C > 0$, which is independent of h. Using this result in (4.42), we get

$$\begin{aligned}\mathrm{e}^{-Ct}\|\boldsymbol{e}(t)\|_{\boldsymbol{\mathcal{X}}} &\le C\|\boldsymbol{\mathcal{J}}y_0 - \boldsymbol{y}_0\|_{\boldsymbol{\mathcal{X}}} + C\int_0^t \mathrm{e}^{-Cs}\|\boldsymbol{e}(s)\|_{\boldsymbol{\mathcal{X}}}\,\mathrm{d}s + tC\sup_{[0,t]}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})y\|_{\boldsymbol{\mathcal{X}}} \\ &\quad + tC(\sup_{[0,t]}\|\mathcal{R}_\Lambda(y)\partial_t y\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t]}\|\mathcal{R}_\mathrm{A} y\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t]}\|\mathcal{R}_\mathrm{F}(\cdot,y)\|_{\boldsymbol{\mathcal{X}}}).\end{aligned}$$

With the Gronwall inequality, we finally get

$$\begin{aligned}\|\boldsymbol{e}(t)\|_{\boldsymbol{\mathcal{X}}} \le C(1+t)\mathrm{e}^{Ct}\Big(&\|\boldsymbol{\mathcal{J}}y_0 - \boldsymbol{y}_0\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t]}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})y\|_{\boldsymbol{\mathcal{X}}}) \\ &+ \sup_{[0,t]}\|\mathcal{R}_\Lambda(y)\partial_t y\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t]}\|\mathcal{R}_\mathrm{A} y\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t]}\|\mathcal{R}_\mathrm{F}(\cdot,y)\|_{\boldsymbol{\mathcal{X}}}\Big),\end{aligned} \tag{4.44}$$

which proves the result. □

In the next step, we combine the wellposedness result from Lemma 4.18 with the approximation property from Theorem 4.20 in order to ensure $t_h^*(\boldsymbol{y}_0) \geq t^*(y_0)$ for h sufficiently small, i.e., provided the discrete approximation is accurate enough, it approximates the continuous solution as long as it exists. To do so, we first define the constant

$$C_{\max}(h) = \max\{1, \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h), \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\boldsymbol{C}_{\mathbf{A}}(h)\}. \tag{4.45}$$

To provide an understanding of the typical behavior of this constant, we indicate the constants appearing in the specific examples.

Example 4.21. *As stated in Section 5.1 below, we obtain for the Westervelt equation*

$$\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h) = Ch^{-\frac{d}{2}}, \qquad \boldsymbol{C}_{\mathbf{A}}(h) = Ch^{-1}, \qquad C_{\max}(h) = Ch^{-1-\frac{d}{2}},$$

where $d \in \mathbb{N}$ is the dimension of the spatial domain Ω. In Section 5.2, we show for the Maxwell equations

$$\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h) = Ch^{-\frac{3}{2}}, \qquad \boldsymbol{C}_{\mathbf{A}}(h) = Ch^{-1}, \qquad C_{\max}(h) = Ch^{-\frac{5}{2}}.$$

In both cases, $C > 0$ is a constant independent of h.

We further assume the following approximation properties based on a space $\mathcal{Z} \hookrightarrow \mathcal{Y}$. For the specific examples, this space is defined in (3.13) and (3.24).

Assumption 4.22. *Let Assumption 4.1 be satisfied and $R > 0$. There exists a Hilbert space $(\mathcal{Z}, (\cdot \mid \cdot)_{\mathcal{Z}})$ such that $\mathcal{Z} \hookrightarrow \mathcal{Y}$ holds with a continuous and dense embedding. Moreover, the space discretization is consistent, i.e., for $h \to 0$ we have*

(A_1) $\|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{J}})\zeta\|_{\mathcal{X}} \to 0$,

(A_2) $C_{\max}(h)\|\boldsymbol{\mathcal{J}}y_0 - \boldsymbol{y}_0\|_{\boldsymbol{\mathcal{X}}} \to 0$,

(A_3) $C_{\max}(h)\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\zeta\|_{\boldsymbol{\mathcal{X}}} \to 0$,

(A_4) $C_{\max}(h)\|\mathcal{R}_{\Lambda}(\xi)\zeta\|_{\boldsymbol{\mathcal{X}}} \to 0$,

(A_5) $C_{\max}(h)\|\mathcal{R}_{\mathrm{A}}\zeta\|_{\boldsymbol{\mathcal{X}}} \to 0$,

(A_6) $C_{\max}(h)\sup_{J_T}\|\mathcal{R}_{\mathrm{F}}(\cdot, \xi)\|_{\boldsymbol{\mathcal{X}}} \to 0$,

uniformly for $\xi, \zeta \in \mathcal{Z}$ with $\xi \in B_{\mathcal{Y}}(R)$.

(A_1) states that the approximation space $\boldsymbol{\mathcal{X}}$, the reference operator $\boldsymbol{\mathcal{J}}$, and the lift operator $\boldsymbol{\mathcal{L}}$ are suitably chosen. (A_2) implies that the discrete initial value $\boldsymbol{y}_0$ converges to its continuous counterpart. Furthermore, (A_3) states that $\boldsymbol{\mathcal{I}}$ and $\boldsymbol{\mathcal{J}}$ are compatible. Finally, the last three assumptions (A_4), (A_5), and (A_6) concern the remainder terms, i.e., these assumptions imply that the discrete operators $\boldsymbol{\Lambda}$, $\mathbf{A}$, and $\mathbf{F}$ are compatible to their continuous counterparts Λ, A, and F, respectively.

We now fix the radii used in the previous results for the rest of this thesis. To do so, we use that Assumption 3.1 and Assumption 4.1 stay true if we reduce the corresponding radii. Thus, we first choose $\boldsymbol{R} > 0$ such that Assumption 4.1 is satisfied. Then, in Assumption 3.1 we choose $R > 0$ sufficiently small such that (4.32) holds. We emphasize that R may not be the optimal radius in Assumption 3.1. Based on this choice, we state the following assumption, which is a sharper version of Assumption 3.2.

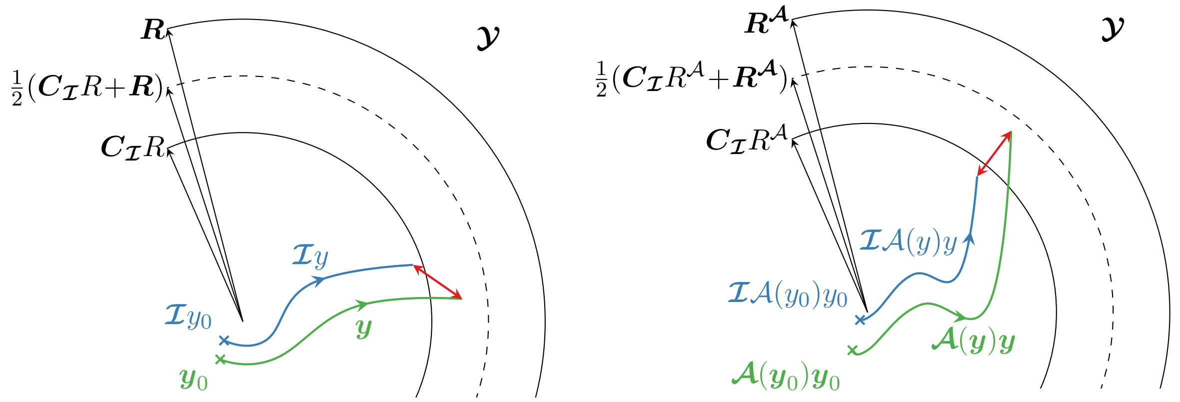

Figure 4.3: Illustration of the different radii used for the solution (left) and the differential operator applied to the solution (right).

Assumption 4.23. *Let Assumption 3.1 be satisfied. The quasilinear Cauchy problem (3.3) has a unique solution with maximal time of existence $t^*(y_0) > 0$ such that for every $T < t^*(y_0)$ there is unique a solution y of (3.3) satisfying*

$$y \in C^1(J_T, \mathcal{Z}) \cap C(J_T, B_{\mathcal{Y}}(R)).$$

Additionally, there are $R^{\partial_t}, R^{\mathcal{A}} > 0$ such that the solution satisfies the bounds

$$\|\partial_t y(t)\|_{\mathcal{Y}} < R^{\partial_t}, \qquad \|\mathcal{A}(y(t))y(t)\|_{\mathcal{Y}} < R^{\mathcal{A}}$$

uniformly for $t \in J_T$.

Finally, we choose $\boldsymbol{R^{\mathcal{A}}} > 0$ with $\boldsymbol{R^{\mathcal{A}}} > \boldsymbol{C_{\mathcal{I}}} R^{\mathcal{A}}$.

As depicted on the left-hand side of Figure 4.3 in blue, Assumption 4.23 yields a maximal time of existence $t^*(y_0)$ such that for all $T < t^*(y_0)$ the solution y of (3.3) satisfies

$$\|\boldsymbol{\mathcal{I}} y(t)\|_{\boldsymbol{\mathcal{Y}}} \le \boldsymbol{C_{\mathcal{I}}} \|y(t)\|_{\mathcal{Y}} < \boldsymbol{C_{\mathcal{I}}} R, \qquad t \in J_T,$$

due to the boundedness (4.18) of $\boldsymbol{\mathcal{I}}$. Based on Lemma 4.18, we further get under suitable assumptions on the initial value $\boldsymbol{y}_0$ the existence of a solution $\boldsymbol{y}$ (green) of the discrete problem (4.12) with maximal time of existence $t_h^*(\boldsymbol{y}_0) > 0$, i.e., we have for all $T_h < t_h^*(\boldsymbol{y}_0)$

$$\|\boldsymbol{y}(t)\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}, \qquad t \in J_{T_h}.$$

Finally, in Theorem 4.20 we proved an error estimate, which in combination with Assumption 4.22 allows us to bound the difference $\boldsymbol{\mathcal{I}} y - \boldsymbol{y}$, which is indicated by the red line.

However, to complete the analysis of the space discretization, we have to prove $t_h^*(\boldsymbol{y}_0) \ge t^*(y_0)$ such that the discrete quasilinear Cauchy problem (4.12) actually yields approximations to the solution of the continuous problem (3.3). This is done in Theorem 4.25, where we use the error estimate to show that

$$\|\boldsymbol{y}\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{C_{\mathcal{I}}} R + \boldsymbol{R})$$

holds for h sufficiently small. This yields the desired result.

Note that based on the radii $R^{\mathcal{A}}$ and $\boldsymbol{R}^{\boldsymbol{\mathcal{A}}}$, we use similar bounds for $\mathcal{A}(y)y$ and $\boldsymbol{\mathcal{A}}(\boldsymbol{y})\boldsymbol{y}$ as for y and $\boldsymbol{y}$, respectively, as illustrated on the right-hand side of Figure 4.3.

As a preparation for the final theorem of this section, we need the following lemma, which is also useful for the analysis of fully discrete schemes.

Lemma 4.24. *Let $\xi \in \mathcal{Z} \cap B_{\mathcal{Y}}(R)$ with $\mathcal{A}(\xi)\xi \in B_{\mathcal{Y}}(R^{\mathcal{A}})$. Further, let $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in \boldsymbol{\mathcal{Y}}$ such that*

$$\|\boldsymbol{\xi}_i\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}, \qquad i = 1, 2,$$

with $\boldsymbol{R}$ independent of h as well as

$$C_{\max}(h)\|\boldsymbol{\mathcal{J}}\xi - \boldsymbol{\xi}_i\|_{\boldsymbol{\mathcal{X}}} \leq C_{\mathrm{conv}}(h), \qquad i = 1, 2, \tag{4.46}$$

where $C_{\mathrm{conv}}(h)$ depends on R, but is independent of ξ. If $C_{\mathrm{conv}}(h) \to 0$ holds for $h \to 0$ and Assumption 4.22 is satisfied, then there exists $h_0 > 0$ such that

$$\|\boldsymbol{\xi}_1\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R), \qquad \|\boldsymbol{\mathcal{A}}(\boldsymbol{\xi}_2)\boldsymbol{\xi}_1\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R^{\mathcal{A}})$$

holds for all $h < h_0$.

Proof. First, we see with (4.17) and (4.1) that

$$\begin{aligned}\|\boldsymbol{\xi}_1\|_{\boldsymbol{\mathcal{Y}}} &\leq \|\boldsymbol{\xi}_1 - \boldsymbol{\mathcal{J}}\xi\|_{\boldsymbol{\mathcal{Y}}} + \|(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{I}})\xi\|_{\boldsymbol{\mathcal{Y}}} + \|\boldsymbol{\mathcal{I}}\xi\|_{\boldsymbol{\mathcal{Y}}} \\ &\leq \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\|\boldsymbol{\xi}_1 - \boldsymbol{\mathcal{J}}\xi\|_{\boldsymbol{\mathcal{X}}} + \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\|(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{I}})\xi\|_{\boldsymbol{\mathcal{X}}} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R\end{aligned}$$

holds. Due to (4.46) and Assumption 4.22, we have for $h \to 0$

$$\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\|\boldsymbol{\xi}_1 - \boldsymbol{\mathcal{J}}\xi\|_{\boldsymbol{\mathcal{X}}} + \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\|(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{I}})\xi\|_{\boldsymbol{\mathcal{X}}} \to 0.$$

Therefore, there exists $h_1 > 0$ such that

$$\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\|\boldsymbol{\xi}_1 - \boldsymbol{\mathcal{J}}\xi\|_{\boldsymbol{\mathcal{X}}} + \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\|(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{I}})\xi\|_{\boldsymbol{\mathcal{X}}} < \tfrac{1}{2}(\boldsymbol{R} - \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R)$$

and hence

$$\|\boldsymbol{\xi}_1\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R)$$

holds for all $h < h_1$. We further derive

$$\boldsymbol{\mathcal{A}}(\boldsymbol{\xi}_2)\boldsymbol{\xi}_1 = \boldsymbol{\mathcal{A}}(\boldsymbol{\xi}_2)(\boldsymbol{\xi}_1 - \boldsymbol{\mathcal{J}}\xi) + (\boldsymbol{\mathcal{A}}(\boldsymbol{\xi}_2) - \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}\xi))\boldsymbol{\mathcal{J}}\xi + \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}\xi)\boldsymbol{\mathcal{J}}\xi. \tag{4.47}$$

For the first term, we get from the inverse estimate (4.1) and the bounds (4.16) and (4.29) for $\boldsymbol{\Lambda}^{-1}$ and $\mathbf{A}$, respectively,

$$\begin{aligned}\|\boldsymbol{\mathcal{A}}(\boldsymbol{\xi}_2)(\boldsymbol{\xi}_1 - \boldsymbol{\mathcal{J}}\xi)\|_{\boldsymbol{\mathcal{Y}}} &\leq \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}\boldsymbol{C}_{\mathbf{A}}(h)\|\boldsymbol{\xi}_1 - \boldsymbol{\mathcal{J}}\xi\|_{\boldsymbol{\mathcal{X}}} \\ &\leq \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}C_{\mathrm{conv}}(h),\end{aligned}$$

where we used (4.46) in the last step. Furthermore, we obtain for the second term from the Lipschitz continuity (4.14) of $\boldsymbol{\mathcal{A}}$

$$\|(\boldsymbol{\mathcal{A}}(\boldsymbol{\xi}_2) - \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}\xi))\boldsymbol{\mathcal{J}}\xi\|_{\boldsymbol{\mathcal{Y}}} \leq \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\boldsymbol{L}_{\boldsymbol{\mathcal{A}}}\|\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}\xi)\boldsymbol{\mathcal{J}}\xi\|_{\boldsymbol{\mathcal{Y}}}(\|\boldsymbol{\xi}_2 - \boldsymbol{\mathcal{J}}\xi\|_{\boldsymbol{\mathcal{X}}} + \|(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{I}})\xi\|_{\boldsymbol{\mathcal{X}}}).$$

Based on (4.46), we get the existence of a constant $C_0 > 0$ independent of h such that

$$\|(\boldsymbol{\mathcal{A}}(\boldsymbol{\xi}_2) - \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}\xi))\boldsymbol{\mathcal{J}}\xi\|_{\boldsymbol{\mathcal{Y}}} \le C_0 \|\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}\xi)\boldsymbol{\mathcal{J}}\xi\|_{\boldsymbol{\mathcal{Y}}}$$

holds for all $h < h_1$. Using these estimates in (4.47) implies

$$\|\boldsymbol{\mathcal{A}}(\boldsymbol{\xi}_2)\boldsymbol{\xi}_1\|_{\boldsymbol{\mathcal{Y}}} \le \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1} C_{\mathrm{conv}}(h) + (C_0 + 1)\|\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}\xi)\boldsymbol{\mathcal{J}}\xi\|_{\boldsymbol{\mathcal{Y}}}. \tag{4.48}$$

We further have

$$\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}\xi)\boldsymbol{\mathcal{J}}\xi = (\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}\xi)\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_\Lambda^*[\xi]\mathcal{A}(\xi))\xi + (\boldsymbol{\mathcal{L}}_\Lambda^*[\xi] - \boldsymbol{\mathcal{J}})\mathcal{A}(\xi)\xi + (\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{I}})\mathcal{A}(\xi)\xi + \boldsymbol{\mathcal{I}}\mathcal{A}(\xi)\xi, \tag{4.49}$$

where we again consider each term separately. For the first term, we obtain with (4.1), the norm equivalence (4.10), and the bound (4.37) the estimate

$$\|(\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}\xi)\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_\Lambda^*[\xi]\mathcal{A}(\xi))\xi\|_{\boldsymbol{\mathcal{Y}}} \le C\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\boldsymbol{c}_{\boldsymbol{\Lambda}}^{-\frac{1}{2}}\|\mathcal{R}_{\mathrm{A}}\xi\|_{\boldsymbol{\mathcal{X}}}.$$

We further employ (4.1), (4.10), and the bound (4.36) to get

$$\|(\boldsymbol{\mathcal{L}}_\Lambda^*[\xi] - \boldsymbol{\mathcal{J}})\mathcal{A}(\xi)\xi\|_{\boldsymbol{\mathcal{Y}}} \le C\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\boldsymbol{c}_{\boldsymbol{\Lambda}}^{-\frac{1}{2}}\|\mathcal{R}_\Lambda(\xi)\mathcal{A}(\xi)\xi\|_{\boldsymbol{\mathcal{X}}}.$$

Finally, (4.1) yields

$$\|(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{I}})\mathcal{A}(\xi)\xi\|_{\boldsymbol{\mathcal{Y}}} \le \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\|(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{I}})\mathcal{A}(\xi)\xi\|_{\boldsymbol{\mathcal{X}}}.$$

Using these bounds in (4.48) and (4.49), we hence obtain

$$\begin{aligned}\|\boldsymbol{\mathcal{A}}(\boldsymbol{\xi}_2)\boldsymbol{\xi}_1\|_{\boldsymbol{\mathcal{Y}}} \le C C_{\mathrm{conv}}(h) + C C_{\max}(h)(\|\mathcal{R}_{\mathrm{A}}\xi\|_{\boldsymbol{\mathcal{X}}} + \|\mathcal{R}_\Lambda(\xi)\mathcal{A}(\xi)\xi\|_{\boldsymbol{\mathcal{X}}} + \|(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{I}})\mathcal{A}(\xi)\xi\|_{\boldsymbol{\mathcal{X}}})\\ + \|\boldsymbol{\mathcal{I}}\mathcal{A}(\xi)\xi\|_{\boldsymbol{\mathcal{Y}}}.\end{aligned}$$

Thus, Assumption 4.22 and (4.46) imply the existence of $h_2 > 0$ such that

$$\begin{aligned}\|\boldsymbol{\mathcal{A}}(\boldsymbol{\xi}_2)\boldsymbol{\xi}_1\|_{\boldsymbol{\mathcal{Y}}} &< \|\boldsymbol{\mathcal{I}}\mathcal{A}(\xi)\xi\|_{\boldsymbol{\mathcal{Y}}} + \tfrac{1}{2}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} - \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R^{\mathcal{A}})\\ &\le \tfrac{1}{2}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R^{\mathcal{A}})\end{aligned}$$

holds for all $h < h_2$, where we employed (4.18) and (4.1) to bound the first term. The result then follows with $h_0 = \min\{h_1, h_2\}$. □

Based on the previous lemma, we now close the proof of the wellposedness of (4.12).

Theorem 4.25. *Let the assumptions of Theorem 4.20, Assumption 4.22, and Assumption 4.23 be satisfied. Then, we have for $h \to 0$*

$$\|y(t) - \boldsymbol{\mathcal{L}}\boldsymbol{y}(t)\|_{\mathcal{X}} \to 0, \qquad t \in [0, \min\{t^*(y_0), t_h^*(\boldsymbol{y}_0)\}). \tag{4.50}$$

Furthermore, there exists $h_0 > 0$ with $t_h^*(\boldsymbol{y}_0) \ge t^*(y_0)$ for all $h < h_0$. In particular, for all $T < t^*(y_0)$ there exists a unique solution $\boldsymbol{y}$ of (4.12) such that (4.30) and (4.31) hold.

Proof. Note that (4.50) follows directly from Assumption 4.22 and the error estimate (4.39), as all terms on the right-hand side tend to zero.

We prove the lower bound on the maximal time of existence $t_h^*(\boldsymbol{y}_0)$ of the discrete solution by contradiction, i.e., we assume that $t_h^*(\boldsymbol{y}_0) < t^*(y_0)$ holds for all $h > 0$, which in particular implies $t_h^*(\boldsymbol{y}_0) < \infty$. As $t_h^*(\boldsymbol{y}_0)$ is the maximal time of existence of the solution $\boldsymbol{y}$ of (4.12), we have

$$\lim_{t \to t_h^*(\boldsymbol{y}_0)} \|\boldsymbol{y}(t)\|_{\boldsymbol{\mathcal{Y}}} = \boldsymbol{R} \qquad \text{or} \qquad \lim_{t \to t_h^*(\boldsymbol{y}_0)} \|\boldsymbol{\mathcal{A}}(\boldsymbol{y}(t))\boldsymbol{y}(t)\|_{\boldsymbol{\mathcal{Y}}} = \boldsymbol{R}^{\boldsymbol{\mathcal{A}}}. \tag{4.51}$$

We now show that there exists $h_0 > 0$ such that (4.51) is false for all $h < h_0$. Let $T < t_h^*(\boldsymbol{y})$ arbitrary. Then, the bound (4.44) for the discrete error together with Assumption 4.22 yields that (4.46) is satisfied with $\xi = y(t)$ and $\boldsymbol{\xi}_i = \boldsymbol{y}(t)$, $i = 1, 2$. Hence, Lemma 4.24 shows the existence of $h_0 > 0$ with

$$\sup_{J_T} \|\boldsymbol{y}(t)\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R} + \boldsymbol{C_{\mathcal{I}}} R), \qquad \sup_{J_T} \|\boldsymbol{\mathcal{A}}(\boldsymbol{y})\boldsymbol{y}\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \boldsymbol{C_{\mathcal{I}}} R^{\mathcal{A}}), \tag{4.52}$$

for $h < h_0$. Furthermore, as both $\boldsymbol{y}$ and $\boldsymbol{\mathcal{A}}(\boldsymbol{y})\boldsymbol{y}$ are continuous in time and the bounds (4.52) are independent of T, we finally get

$$\lim_{t \to t_h^*(\boldsymbol{y}_0)} \|\boldsymbol{y}(t)\|_{\boldsymbol{\mathcal{Y}}} \leq \tfrac{1}{2}(\boldsymbol{R} + \boldsymbol{C_{\mathcal{I}}} R) \quad \text{and} \quad \lim_{t \to t_h^*(\boldsymbol{y}_0)} \|\boldsymbol{\mathcal{A}}(\boldsymbol{y}(t))\boldsymbol{y}(t)\|_{\boldsymbol{\mathcal{Y}}} \leq \tfrac{1}{2}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \boldsymbol{C_{\mathcal{I}}} R^{\mathcal{A}})$$

for all $h < h_0$, which is a contradiction to (4.51), since $\boldsymbol{C_{\mathcal{I}}} R < \boldsymbol{R}$ and $\boldsymbol{C_{\mathcal{I}}} R^{\mathcal{A}} < \boldsymbol{R}^{\boldsymbol{\mathcal{A}}}$. □

In the next section, we will discuss a suitable discretization of the nonlinearities Λ and F satisfying Assumption 4.1.

4.4 Discretization of local nonlinearities

We now investigate the remainder terms $\mathcal{R}_\Lambda$ and $\mathcal{R}_{\mathrm{F}}$. In order to derive bounds in terms of the operators introduced at the end of Section 4.1, we provide a specific choice for the discretization of the nonlinearities Λ and F.

To do so, we narrow the abstract framework presented in the previous section down to the space discretization of partial differential equations. In particular, for some $d, d_r \in \mathbb{N}$ the spaces $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ are function spaces from a bounded domain $\Omega \subset \mathbb{R}^d$ to $\mathbb{R}^{d_r}$. Correspondingly, the discrete spaces $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$ are function spaces from a bounded domain $\Omega_h \subset \mathbb{R}^d$ to $\mathbb{R}^{d_r}$.

Based on these spaces, we require the following additional assumption, which states that the nonlinearities are not only local in time but also local in space.

Assumption 4.26. *We have the following properties of the nonlinearities as well as the interpolation and lift operator.*

(λ) *We have $\Lambda(\xi) \in \mathcal{L}(\mathcal{Y})$ for $\xi \in B_{\mathcal{Y}}(R)$. Furthermore, the operator Λ is local in space, i.e., there exists a map $\lambda : \Omega \times \mathbb{R}^{d_r} \to \mathbb{R}^{d_r \times d_r}$ such that for every $\xi \in B_{\mathcal{Y}}(R)$ and $\varphi \in \mathcal{X}$ the identity*

$$(\Lambda(\xi)\varphi)(x) = \lambda(x, \xi(x))\varphi(x), \qquad x \in \Omega, \tag{4.53}$$

holds in $\mathcal{X}$.

(f) *The nonlinearity* F *is local in space, i.e., there exists a map* $f : J_T \times \Omega \times \mathbb{R}^{d_r} \to \mathbb{R}^{d_r}$ *such that for every* $\xi \in B_{\mathcal{Y}}(R)$ *and* $t \in J_T$*, the identity*

$$(\mathrm{F}(t,\xi))(x) = f(t,x,\xi(x)), \qquad x \in \Omega, \tag{4.54}$$

is satisfied in $\mathcal{X}$.

($\boldsymbol{\mathcal{I}}, \boldsymbol{\mathcal{L}}$) *The operator* $\boldsymbol{\mathcal{I}}$ *is a nodal interpolation operator, i.e., there exist* $M \in \mathbb{N}$*, interpolation points* $\Omega_{\boldsymbol{\mathcal{I}}} = \{\underline{x}_0, \dots, \underline{x}_M\} \subset \Omega \cap \Omega_h$*, and basis functions* $\{\boldsymbol{\phi}_0, \dots, \boldsymbol{\phi}_M\} \subset \boldsymbol{\mathcal{Y}}$ *such that*

$$\boldsymbol{\mathcal{I}}\xi = \sum_{m=0}^{M} \xi(\underline{x}_m)\, \boldsymbol{\phi}_m, \qquad \boldsymbol{\mathcal{I}}\xi(\underline{x}) = \xi(\underline{x}), \qquad \xi \in \mathcal{Y},\ \underline{x} \in \Omega_{\boldsymbol{\mathcal{I}}},$$

holds. Further, the lift operator $\boldsymbol{\mathcal{L}}$ *preserves the values at the interpolation points, i.e., we have*

$$\boldsymbol{\mathcal{L}}\boldsymbol{\xi}(\underline{x}) = \boldsymbol{\xi}(\underline{x}), \qquad \boldsymbol{\xi} \in \boldsymbol{\mathcal{X}},\ \underline{x} \in \Omega_{\boldsymbol{\mathcal{I}}}.$$

Assumption 4.26 allows us to define the discrete operator $\boldsymbol{\Lambda}$ corresponding to Λ by

$$\boldsymbol{\Lambda}(\boldsymbol{\xi})\boldsymbol{\varphi} := \boldsymbol{\mathcal{I}}\Lambda(\boldsymbol{\mathcal{L}}\boldsymbol{\xi})\boldsymbol{\mathcal{L}}\boldsymbol{\varphi} \tag{4.55a}$$

$$= \sum_{m=0}^{M} \lambda(\underline{x}_m, \boldsymbol{\xi}(\underline{x}_m))\boldsymbol{\varphi}(\underline{x}_m)\, \boldsymbol{\phi}_m, \tag{4.55b}$$

for $\boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$ and $\boldsymbol{\psi} \in \boldsymbol{\mathcal{X}}$. Note that, despite (4.55a) is motivated by (4.53), we have to be cautious with this notation, as it does not yield a well-defined operator on its own, i.e., we do not have $\boldsymbol{\mathcal{L}}\boldsymbol{\xi} \in B_{\mathcal{Y}}(R)$ in general. We also get $\Lambda(\boldsymbol{\mathcal{L}}\boldsymbol{\xi})\boldsymbol{\mathcal{L}}\boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}}$, but $\boldsymbol{\mathcal{I}}$ is only defined on $\mathcal{Y}$. Nevertheless, this expression is well defined in the sense of (4.55b), since $\boldsymbol{\mathcal{I}}$ and $\boldsymbol{\mathcal{L}}$ preserve the interpolation points.

Hence, we use (4.55a) in the following for the sake of readability, keeping (4.55b) in mind. We point out that the notation (4.55a) makes sense, as the lift operator $\boldsymbol{\mathcal{L}}$ maps functions from the discrete to the continuous space such that the continuous operator Λ can be applied. The interpolation operator $\boldsymbol{\mathcal{I}}$ then maps the continuous result back to the discrete space.

For the nonlinearity $\mathbf{F}$, we follow a similar approach, i.e., we define

$$\mathbf{F}(t,\boldsymbol{\xi}) := \boldsymbol{\mathcal{I}}\mathrm{F}(t,\boldsymbol{\mathcal{L}}\boldsymbol{\xi}) \tag{4.56a}$$

$$= \sum_{m=0}^{M} f(t, \underline{x}_m, \boldsymbol{\xi}(\underline{x}_m))\, \boldsymbol{\phi}_m, \tag{4.56b}$$

for $t \in J_T$ and $\boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$. Note that (4.56a) is motivated by (4.54), but again only well defined in the sense of (4.56b).

As it is not possible to verify the properties of $\boldsymbol{\Lambda}$ and $\mathbf{F}$ in this general framework, we just assume that Assumption 4.1 is satisfied by $\boldsymbol{\Lambda}$ and $\mathbf{F}$ defined by (4.55) and (4.56), respectively. Note that we prove this for the specific examples in Sections 5.1 and 5.2.

Based on this assumption and the definition

$$\triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\zeta}) := \sup_{\|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}}=1} \Big((\boldsymbol{\zeta} \mid \boldsymbol{\xi})_{\boldsymbol{\mathcal{X}}} - (\boldsymbol{\mathcal{L}}\boldsymbol{\zeta} \mid \boldsymbol{\mathcal{L}}\boldsymbol{\xi})_{\mathcal{X}} \Big), \qquad \boldsymbol{\zeta}, \boldsymbol{\xi} \in \boldsymbol{\mathcal{X}}, \tag{4.57}$$

we now derive estimates for the remainder terms.

Lemma 4.27. *If Assumption 4.26 is satisfied, then we have for $t \in J_T$, $\xi \in B_{\mathcal{Y}}(R^{\partial_t})$, and $\zeta \in B_{\mathcal{Y}}(R)$ the bounds*

$$\|\mathcal{R}_\Lambda(\zeta)\xi\|_{\boldsymbol{\mathcal{X}}} \le \boldsymbol{C}_{\boldsymbol{\Lambda}}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\xi\|_{\boldsymbol{\mathcal{X}}} + \boldsymbol{C}_{\boldsymbol{\mathcal{L}}}\|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}})\Lambda(\zeta)\xi\|_{\mathcal{X}} + \triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}\Lambda(\zeta)\xi)\,, \tag{4.58}$$
$$\|\mathcal{R}_{\mathrm{F}}(t,\zeta)\|_{\boldsymbol{\mathcal{X}}} \le \boldsymbol{C}_{\boldsymbol{\mathcal{L}}}\|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}})\mathrm{F}(t,\zeta)\|_{\mathcal{X}} + \triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}\mathrm{F}(t,\zeta))\,. \tag{4.59}$$

Proof. Let $t \in J_T$, $\xi \in B_{\mathcal{Y}}(R^{\partial_t})$, and $\zeta \in B_{\mathcal{Y}}(R)$ arbitrary. The definition (4.33) of $\mathcal{R}_\Lambda$ together with the definition (4.55) of $\boldsymbol{\Lambda}$ implies

$$\begin{aligned}\mathcal{R}_\Lambda(\zeta) &= \boldsymbol{\mathcal{I}}\Lambda(\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}}\zeta)\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_{\mathcal{X}}^*\Lambda(\zeta)\\ &= \boldsymbol{\mathcal{I}}\Lambda(\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}}\zeta)\boldsymbol{\mathcal{L}}(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{I}}) + \boldsymbol{\mathcal{I}}\Lambda(\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}}\zeta)(\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}} - \mathrm{Id}) + \boldsymbol{\mathcal{I}}(\Lambda(\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}}\zeta) - \Lambda(\zeta)) + (\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{L}}_{\mathcal{X}}^*)\Lambda(\zeta).\end{aligned}$$

Using that the interpolation points are preserved by $\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}}$, we see that both intermediate terms vanish, i.e., we have

$$\boldsymbol{\mathcal{I}}\Lambda(\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}}\zeta)(\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}} - \mathrm{Id})\xi = \sum_{m=0}^{M} \lambda(\underline{x}_m, \zeta(\underline{x}_m))(\xi(\underline{x}_m) - \xi(\underline{x}_m))\,\boldsymbol{\phi}_m = 0,$$

and

$$\boldsymbol{\mathcal{I}}(\Lambda(\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}}\zeta) - \Lambda(\zeta))\xi = \sum_{m=0}^{M} (\lambda(\underline{x}_m, \zeta(\underline{x}_m)) - \lambda(\underline{x}_m, \zeta(\underline{x}_m)))\xi(\underline{x}_m)\,\boldsymbol{\phi}_m = 0.$$

We further get from (4.3) the bound

$$\|\boldsymbol{\mathcal{I}}\Lambda(\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}}\zeta)\boldsymbol{\mathcal{L}}(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{I}})\xi\|_{\boldsymbol{\mathcal{X}}} = \|\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\zeta)(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{I}})\xi\|_{\boldsymbol{\mathcal{X}}} \le \boldsymbol{C}_{\boldsymbol{\Lambda}}\|(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{I}})\xi\|_{\boldsymbol{\mathcal{X}}}.$$

The definition (4.20) of the adjoint lift operator $\boldsymbol{\mathcal{L}}_{\mathcal{X}}^*$ yields

$$\begin{aligned}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{L}}_{\mathcal{X}}^*)\Lambda(\zeta)\xi\|_{\boldsymbol{\mathcal{X}}} &= \sup_{\|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}}=1} ((\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{L}}_{\mathcal{X}}^*)\Lambda(\zeta)\xi \mid \boldsymbol{\xi})_{\boldsymbol{\mathcal{X}}}\\ &= \sup_{\|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}}=1} \Big((\boldsymbol{\mathcal{I}}\Lambda(\zeta)\xi \mid \boldsymbol{\xi})_{\boldsymbol{\mathcal{X}}} - (\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}}\Lambda(\zeta)\xi \mid \boldsymbol{\mathcal{L}}\boldsymbol{\xi})_{\mathcal{X}}\\ &\qquad + ((\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}} - \mathrm{Id})\Lambda(\zeta)\xi \mid \boldsymbol{\mathcal{L}}\boldsymbol{\xi})_{\mathcal{X}}\Big).\end{aligned} \tag{4.60}$$

Finally, the boundedness (4.19) of the lift operator $\boldsymbol{\mathcal{L}}$ yields

$$\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{L}}_{\mathcal{X}}^*)\Lambda(\zeta)\xi\|_{\boldsymbol{\mathcal{X}}} \le \triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}\Lambda(\zeta)\xi) + \boldsymbol{C}_{\boldsymbol{\mathcal{L}}}\|(\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}} - \mathrm{Id})\Lambda(\zeta)\xi\|_{\mathcal{X}},$$

which proves (4.58).

Similarly, we obtain with the definition (4.35) of the remainder $\mathcal{R}_{\mathrm{F}}$

$$\begin{aligned}\mathcal{R}_{\mathrm{F}}(t,\zeta) &= \boldsymbol{\mathcal{I}}\mathrm{F}(t, \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}}\zeta) - \boldsymbol{\mathcal{L}}_{\mathcal{X}}^*\mathrm{F}(t,\zeta)\\ &= \boldsymbol{\mathcal{I}}(\mathrm{F}(t, \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}}\zeta) - \mathrm{F}(t,\zeta)) + (\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{L}}_{\mathcal{X}}^*)\mathrm{F}(t,\zeta).\end{aligned}$$

As the first term vanishes due to

$$\boldsymbol{\mathcal{I}}(\mathrm{F}(t, \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}}\zeta) - \mathrm{F}(t,\zeta)) = \sum_{m=0}^{M} (f(t, \underline{x}_m, \zeta(\underline{x}_m)) - f(t, \underline{x}_m, \zeta(\underline{x}_m)))\xi(\underline{x}_m)\,\boldsymbol{\phi}_m = 0,$$

and the second term can be bounded as in (4.60), this yields (4.59). □

Finally, we use these estimates to refine the abstract error bound from Theorem 4.20 based on the following assumption, which is a refined version of Assumption 4.22.

Assumption 4.28. *Let Assumption 4.1 be satisfied, $R > 0$ and $C_{\max}(h)$ as defined in (4.45). There exists a Hilbert space $(\mathcal{Z}, (\cdot \mid \cdot)_{\mathcal{Z}})$ such that $\mathcal{Z} \hookrightarrow \mathcal{Y}$ holds with a continuous and dense embedding. Moreover, the space discretization is consistent, i.e., for $h \to 0$ we have*

(A_1) $\|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{J}})\zeta\|_{\mathcal{X}} \to 0,$

(A_2) $C_{\max}(h)\|\boldsymbol{\mathcal{J}} y_0 - \boldsymbol{y}_0\|_{\boldsymbol{\mathcal{X}}} \to 0,$

(A_3) $C_{\max}(h)\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\zeta\|_{\boldsymbol{\mathcal{X}}} \to 0,$

(A_5) $C_{\max}(h)\|\mathcal{R}_{\mathrm{A}}\zeta\|_{\boldsymbol{\mathcal{X}}} \to 0,$

(A_7) $\triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\xi}) \to 0,$

uniformly for $\xi, \zeta \in \mathcal{Z}$ with $\xi \in B_{\mathcal{Y}}(R)$ and $\boldsymbol{\xi} \in \boldsymbol{\mathcal{X}}$.

To conclude this section, we state the following refined version of the wellposedness result Theorem 4.25 for the space discretization.

Corollary 4.29. *Let Assumption 4.26 be satisfied. Then, the statement of Theorem 4.25 is also valid if we replace Assumption 4.22 by Assumption 4.28. In this case, the error satisfies for $T < \min\{t^*(y_0), t_h^*(\boldsymbol{y}_0)\}$ the estimate*

$$\begin{aligned} \|y(t) - \boldsymbol{\mathcal{L}}\boldsymbol{y}(t)\|_{\mathcal{X}} \le \|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{J}})y(t)\|_{\mathcal{X}} + C(1+t)\mathrm{e}^{Ct}\Big(\|\boldsymbol{\mathcal{J}} y_0 - \boldsymbol{y}_0\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t]}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})y\|_{\boldsymbol{\mathcal{X}}} \\ + \sup_{[0,t]}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\partial_t y\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t]}\|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}})\Lambda(y)\partial_t y\|_{\mathcal{X}} + \sup_{[0,t]} \triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}\Lambda(y)\partial_t y) \\ + \sup_{[0,t]}\|\mathcal{R}_{\mathrm{A}} y\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t]}\|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}})\mathrm{F}(\cdot, y)\|_{\mathcal{X}} + \sup_{[0,t]} \triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}\mathrm{F}(\cdot, y))\Big), \end{aligned} \tag{4.61}$$

with a constant $C > 0$, which is independent of h.

CHAPTER 5

Space discretization of the specific examples

In this chapter we apply the abstract results for the space discretization of quasilinear wave-type problems presented in Chapter 4 to the specific examples from Chapter 3.

5.1 Example: Westervelt equation

In this section, we apply the abstract theory for the space discretization of quasilinear wave-type equations from Section 4.3 to the Westervelt equation presented in Section 3.2. Hence, we first introduce the approximation spaces and operators needed. Next, we prove that the discrete operators from Section 4.4 satisfy Assumption 4.1. Finally, since Assumption 4.28 holds, we prove the error estimate for the space discretization based on Corollary 4.29.

For the case $d = 1$, i.e., where the spatial domain $\Omega \subset \mathbb{R}$ is an interval, the boundary of Ω contains only two points. Hence, we can use a conforming space discretization. Furthermore, Sobolev's embedding (2.5) implies for $d = 1$ and $\mathcal{Y}$ given by (3.9) the relation $\mathcal{Y} \subset H_0^1(\Omega)^2 \hookrightarrow C(\Omega)^2$, which is essential to define the nodal interpolation operator.

However, for the case $d \in \{2, 3\}$, we have to use a non-conforming space discretization, as the boundary is assumed to be regular. Furthermore, we use the refined definition (3.12) of $\mathcal{Y}$ to ensure the embedding of $\mathcal{Y}$ into the continuous functions.

Hence, as these two cases differ substantially, we treat them separately.

5.1.1 Example: Westervelt equation (1D)

As discussed above, we start with the case $d = 1$, i.e., where the spatial domain $\Omega = (\omega_-, \omega_+) \subset \mathbb{R}$ is an interval. As the first step, we introduce approximation spaces and define the discrete operators. Next, we prove that these operators satisfy Assumption 4.1. Finally, we apply the abstract theory from Section 4.3 and Section 4.4 to derive an error estimate.

Approximation spaces

In the following, we define the discrete spaces $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$. Based on these spaces, we then specify the operators introduced at the end of Section 4.1 as well as the discrete operators appearing in the discrete quasilinear problem (4.12). This is based on [Ern and Guermond, 2004, Chap. 1.1].

For the discretization, we use Lagrange elements of order $p \in \mathbb{N}$. More precisely, we consider a family $\{\mathcal{T}_h\}_{h>0}$ of partitions of Ω into distinct subintervals. Each partition $\mathcal{T}_h$ then corresponds to a number $L_h \in \mathbb{N}$ and a set of points $\{\omega_{h,0}, \dots, \omega_{h,L+1}\} \subset \mathbb{R}$ with

$$\omega_- = \omega_{h,0} < \omega_{h,1} < \dots < \omega_{h,L} < \omega_{h,L+1} = \omega_+, \qquad h = \max_{\ell \leq L} |\omega_{h,\ell+1} - \omega_{h,\ell}|.$$

For the sake of presentation, we drop the subscript h in the following and write L and ω_ℓ instead of L_h and $\omega_{h,\ell}$, respectively.

For $p \in \mathbb{N}$ the approximation space for v is then given by

$$\boldsymbol{\mathcal{V}}_{\mathcal{H}} \coloneqq \{\varphi \in C(\Omega) \mid \varphi|_{(\omega_\ell,\omega_{\ell+1})} \in \mathcal{P}^p\big((\omega_\ell, \omega_{\ell+1})\big)\} \subset H^1(\Omega),$$

which is the space of continuous functions that are piecewise polynomial with degree at most p on every subinterval $(\omega_\ell, \omega_{\ell+1})$, for $\ell \leq L$. Furthermore, let $M \in \mathbb{N}$ such that there is a consisting of piecewise Lagrange polynomials with corresponding nodes $\{\underline{x}_0, \dots, \underline{x}_M\}$. In particular, with the Kronecker delta δ_{ij} this implies $\boldsymbol{\phi}_i(\underline{x}_j) = \delta_{ij}$. Finally, we define $\boldsymbol{\mathcal{V}}_{\mathcal{V}} \coloneqq \boldsymbol{\mathcal{V}}_{\mathcal{H}} \cap H^1_0(\Omega)$.

Based on these definitions, we define $\boldsymbol{\mathcal{V}} = \boldsymbol{\mathcal{V}}_{\mathcal{V}} \times \boldsymbol{\mathcal{V}}_{\mathcal{H}}$ as the function space for $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{X}}_{\mathcal{V}} \times \boldsymbol{\mathcal{X}}_{\mathcal{H}}$ and $\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{Y}}_{\mathcal{V}} \times \boldsymbol{\mathcal{Y}}_{\mathcal{H}}$. As we have $\boldsymbol{\mathcal{V}} \subset \mathcal{X}$, we further use the same inner product for $\boldsymbol{\mathcal{X}}$ as for $\mathcal{X}$, i.e., we have

$$(\boldsymbol{\varphi} \mid \boldsymbol{\psi})_{\boldsymbol{\mathcal{X}}} = (\boldsymbol{\varphi}_{\mathcal{V}} \mid \boldsymbol{\psi}_{\mathcal{V}})_{H^1_0(\Omega)} + (\boldsymbol{\varphi}_{\mathcal{H}} \mid \boldsymbol{\psi}_{\mathcal{H}})_{L^2(\Omega)} = (\boldsymbol{\varphi} \mid \boldsymbol{\psi})_{\mathcal{X}}\,, \quad \boldsymbol{\varphi} = \begin{pmatrix} \boldsymbol{\varphi}_{\mathcal{V}} \\ \boldsymbol{\varphi}_{\mathcal{H}} \end{pmatrix}, \boldsymbol{\psi} = \begin{pmatrix} \boldsymbol{\psi}_{\mathcal{V}} \\ \boldsymbol{\psi}_{\mathcal{H}} \end{pmatrix} \in \boldsymbol{\mathcal{V}}^2. \tag{5.1}$$

For the norm of $\boldsymbol{\mathcal{Y}}$, we set

$$\|\boldsymbol{\xi}\|^2_{\boldsymbol{\mathcal{Y}}} = \|\boldsymbol{\xi}_{\mathcal{V}}\|^2_{L^\infty(\Omega)} + \|\boldsymbol{\xi}_{\mathcal{H}}\|^2_{L^\infty(\Omega)}, \qquad \boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{\xi}_{\mathcal{V}} \\ \boldsymbol{\xi}_{\mathcal{H}} \end{pmatrix} \in \boldsymbol{\mathcal{V}}^2. \tag{5.2}$$

We point out that it would also be possible to consider piecewise Sobolev norms corresponding to the norm of $\mathcal{Y}$ on every subinterval. However, using the norm (5.2) simplifies the estimates.

In order to derive the inverse estimate (4.1), we further assume the family $\{\mathcal{T}_h\}_{h>0}$ of partitions to be quasi-uniform, i.e., there is a constant $C > 0$ independent of h such that

$$Ch \leq \min_{\ell<L} |\omega_{\ell+1} - \omega_\ell|$$

holds for all partitions in $\{\mathcal{T}_h\}_{h>0}$. Then, [Ern and Guermond, 2004, Cor. 1.141] yields for $\boldsymbol{\varphi} \in \boldsymbol{\mathcal{V}}_{\mathcal{V}}$ the bound,

$$\|\boldsymbol{\varphi}\|_{L^\infty(\Omega)} \leq C_{\text{inv}} h^{-\frac{1}{2}} \|\boldsymbol{\varphi}\|_{L^2(\Omega)}, \qquad |\boldsymbol{\varphi}|_{H^1(\Omega)} \leq C_{\text{inv}} h^{-1} \|\boldsymbol{\varphi}\|_{L^2(\Omega)}, \tag{5.3}$$

with a constant $C_{\text{inv}} > 0$ independent of h. Using the Poincaré inequality, this yields the first estimate in (4.1) for $\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h) = Ch^{-\frac{1}{2}}$. Furthermore, as Ω is bounded, this also implies the second estimate in (4.1) for $\boldsymbol{C}_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}}(h) = Ch^{-1} + C$.

Discrete operators

Next, we focus on the definition of the operators. First, we see that as we consider a conforming discretization, we can choose the lift

$$\boldsymbol{\mathcal{L}} = \begin{pmatrix} \boldsymbol{\mathcal{L}}_{\mathcal{V}} & 0 \\ 0 & \boldsymbol{\mathcal{L}}_{\mathcal{H}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{5.4}$$

Therefore, (4.19) is trivially satisfied. Hence, (4.20) yields

$$\boldsymbol{\mathcal{L}}_{\mathcal{X}}^* = \begin{pmatrix} \boldsymbol{\mathcal{L}}_{\mathcal{V}}^* & 0 \\ 0 & \boldsymbol{\mathcal{L}}_{\mathcal{H}}^* \end{pmatrix} = \boldsymbol{\Pi}_{\boldsymbol{\mathcal{X}}} \tag{5.5}$$

where $\boldsymbol{\Pi}_{\boldsymbol{\mathcal{X}}} : \mathcal{X} \to \boldsymbol{\mathcal{X}}$ denotes the $\mathcal{X}$-orthogonal projection, i.e., we have

$$(\boldsymbol{\Pi}_{\boldsymbol{\mathcal{X}}} \varphi \mid \boldsymbol{\varphi})_{\boldsymbol{\mathcal{X}}} = (\varphi \mid \boldsymbol{\varphi})_{\mathcal{X}}, \qquad \varphi \in \mathcal{X}, \, \boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}}.$$

Based on the basis $\{\boldsymbol{\phi}_0, \dots, \boldsymbol{\phi}_M\}$, we furthermore employ the Lagrange interpolation operator $\boldsymbol{\mathcal{I}}_{\mathcal{V}} : C(\overline{\Omega}) \to C(\overline{\Omega})$, which is given by

$$\boldsymbol{\mathcal{I}}_{\mathcal{V}} \xi = \sum_{m=0}^{M} \xi(\underline{x}_m) \boldsymbol{\phi}_m, \qquad \xi \in C(\overline{\Omega}). \tag{5.6}$$

Moreover, there exists a constant $C_{\boldsymbol{\mathcal{I}},\infty} > 0$ such that

$$C_{\boldsymbol{\mathcal{I}},\infty} = \Big\| \sum_{m=0}^{M} |\boldsymbol{\phi}_m| \Big\|_{L^\infty(\Omega)}$$

holds. In particular, this constant is independent of h since the number of non-vanishing basis functions on every subinterval is only depending on the polynomial degree p. Thus, we have

$$\|\boldsymbol{\mathcal{I}}_{\mathcal{V}} \xi\|_{L^\infty(\Omega)} \leq C_{\boldsymbol{\mathcal{I}},\infty} \|\xi\|_{L^\infty(\Omega)}, \qquad \xi \in C(\overline{\Omega}). \tag{5.7}$$

For $\boldsymbol{\mathcal{I}}_{\mathcal{H}} = \boldsymbol{\mathcal{I}}_{\mathcal{V}}$, we hence set

$$\boldsymbol{\mathcal{I}} = \begin{pmatrix} \boldsymbol{\mathcal{I}}_{\mathcal{V}} & 0 \\ 0 & \boldsymbol{\mathcal{I}}_{\mathcal{H}} \end{pmatrix}, \qquad \boldsymbol{\mathcal{J}} = \begin{pmatrix} \boldsymbol{\mathcal{L}}_{\mathcal{V}}^* & 0 \\ 0 & \boldsymbol{\mathcal{I}}_{\mathcal{H}} \end{pmatrix}. \tag{5.8}$$

Note that the boundedness (4.18) of $\boldsymbol{\mathcal{I}}$ follows from Sobolev's embedding (2.5). Hence, the boundedness (4.17) of $\boldsymbol{\mathcal{J}}$ follows from the boundedness (4.22) of the adjoint lift $\boldsymbol{\mathcal{L}}_{\mathcal{X}}^*$.

Furthermore, as the nonlinearities (3.8) appearing in the Westervelt equation are local, we now employ the approach presented in Section 4.4 to derive their discrete counterparts. This yields for some $\boldsymbol{R} > 0$, which is specified below, the discrete nonlinearities

$$\boldsymbol{\Lambda}(\boldsymbol{\xi}) = \begin{pmatrix} 1 & 0 \\ 0 & \boldsymbol{\Lambda}_{\mathcal{H}}(\boldsymbol{\xi}_{\mathcal{V}}) \end{pmatrix}, \qquad \mathbf{F}(\boldsymbol{\xi}) = \begin{pmatrix} 0 \\ \mathbf{F}_{\mathcal{H}}(\boldsymbol{\xi}_{\mathcal{H}}) \end{pmatrix}, \qquad \boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{\xi}_{\mathcal{V}} \\ \boldsymbol{\xi}_{\mathcal{H}} \end{pmatrix} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}), \tag{5.9}$$

where $\boldsymbol{\Lambda}_{\mathcal{H}}(\boldsymbol{\xi}_{\mathcal{V}}) : \boldsymbol{\mathcal{X}}_{\mathcal{H}} \to \boldsymbol{\mathcal{X}}_{\mathcal{H}}$ and $\mathbf{F}_{\mathcal{H}}(\boldsymbol{\xi}_{\mathcal{H}})$ are given by

$$\boldsymbol{\Lambda}_{\mathcal{H}}(\boldsymbol{\xi}_{\mathcal{V}}) \boldsymbol{\zeta} = \boldsymbol{\mathcal{I}}_{\mathcal{H}}((1 - \varkappa \boldsymbol{\xi}_{\mathcal{V}}) \boldsymbol{\zeta}), \qquad \mathbf{F}_{\mathcal{H}}(\boldsymbol{\xi}_{\mathcal{H}}) = \boldsymbol{\mathcal{I}}_{\mathcal{H}}(\varkappa \boldsymbol{\xi}_{\mathcal{H}}^2), \qquad \boldsymbol{\zeta} \in \boldsymbol{\mathcal{X}}_{\mathcal{H}}.$$

As in the linear case, we introduce the discrete Laplacian $\mathbf{\Delta} : \boldsymbol{\mathcal{X}}_{\mathcal{V}} \to \boldsymbol{\mathcal{X}}_{\mathcal{H}}$ by

$$(\mathbf{\Delta}\, \boldsymbol{\varphi} \mid \boldsymbol{\psi})_{L^2(\Omega)} := -\, (\nabla\, \boldsymbol{\varphi} \mid \nabla\, \boldsymbol{\psi})_{L^2(\Omega)}\,, \qquad \boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}}_{\mathcal{V}},\ \boldsymbol{\psi} \in \boldsymbol{\mathcal{X}}_{\mathcal{H}}, \tag{5.10}$$

which yields

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ \mathbf{\Delta} & 0 \end{pmatrix}. \tag{5.11}$$

Finally, we use the interpolation operator to compute discrete approximations

$$\begin{pmatrix} \boldsymbol{u_0} \\ \boldsymbol{v_0} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mathcal{I}}_{\mathcal{V}} u_0 \\ \boldsymbol{\mathcal{I}}_{\mathcal{H}} v_0 \end{pmatrix} \tag{5.12}$$

to the initial values (u_0, v_0). Thus, the discrete Westervelt equation can be written in the form (4.2), i.e., we seek the solution $(\boldsymbol{u}, \boldsymbol{v}) : J_T \to \boldsymbol{\mathcal{X}}$ of

$$\begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{\Lambda}_{\mathcal{H}}(\boldsymbol{u}) \end{pmatrix} \begin{pmatrix} \partial_t \boldsymbol{u} \\ \partial_t \boldsymbol{v} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \mathbf{\Delta} & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix} + \begin{pmatrix} 0 \\ \mathbf{F}_{\mathcal{H}}(\boldsymbol{v}) \end{pmatrix} \quad \text{on } J_T \\ \boldsymbol{u}(0) = \boldsymbol{u_0}, \qquad \boldsymbol{v}(0) = \boldsymbol{v_0}. \end{cases} \tag{5.13}$$

In the next section, we prove that the operators introduced above satisfy Assumption 4.1.

Properties of the discrete operators

In the following, we prove that, for a suitable choice of $\boldsymbol{R}$, the operators defined in the previous section satisfy Assumption 4.1. To do so, we first introduce a discrete norm, which is equivalent to the norm of $\boldsymbol{\mathcal{X}}_{\mathcal{H}}$.

More precisely, we define the discrete norm

$$|||\boldsymbol{\varphi}|||^2 = h \sum_{m=0}^{M} |\boldsymbol{\varphi}(\underline{x}_m)|^2, \qquad \boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}}_{\mathcal{H}}, \tag{5.14}$$

which is equivalent to the norm of $\boldsymbol{\mathcal{X}}_{\mathcal{H}}$, i.e., there are constants $c_{\mathrm{norm}}, C_{\mathrm{norm}} > 0$ uniformly in $h > 0$ such that

$$c_{\mathrm{norm}} \|\boldsymbol{\varphi}\|^2_{\boldsymbol{\mathcal{X}}_{\mathcal{H}}} \leq |||\boldsymbol{\varphi}|||^2 \leq C_{\mathrm{norm}} \|\boldsymbol{\varphi}\|^2_{\boldsymbol{\mathcal{X}}_{\mathcal{H}}}, \qquad \boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}}_{\mathcal{H}}. \tag{5.15}$$

For a detailed proof, we refer to [Leibold, 2017, Lem. 5.2]. Further, due to the definition (5.6) of the nodal interpolation operator, we have

$$|||\boldsymbol{\mathcal{I}}\varphi||| = |||\varphi|||, \qquad \varphi \in \mathcal{Y}_{\mathcal{H}}. \tag{5.16}$$

Using this norm, we investigate the discrete operators in the following lemma.

Lemma 5.1. *Let $\boldsymbol{R} \in (0, \frac{1}{|\varkappa|})$, with $\varkappa$ introduced in (3.6). Then, the discrete operators $\mathbf{\Lambda}$, $\mathbf{A}$, and $\mathbf{F}$ given in (5.9) and (5.11) satisfy Assumption 4.1.*

Proof. ($\mathbf{\Lambda}$) To prove the assumptions on $\mathbf{\Lambda}$, let $\boldsymbol{\xi} = (\boldsymbol{\xi}_{\mathcal{V}}, \boldsymbol{\xi}_{\mathcal{H}}) \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$ and $\boldsymbol{\varphi} = (\boldsymbol{\varphi}_{\mathcal{V}}, \boldsymbol{\varphi}_{\mathcal{H}}) \in \boldsymbol{\mathcal{X}}$. We first use the norm equivalence (5.15) and (5.16) to derive

$$\begin{aligned}\|\mathbf{\Lambda}(\boldsymbol{\xi})\boldsymbol{\varphi}\|_{\boldsymbol{\mathcal{X}}}^2 &= |\boldsymbol{\varphi}_{\mathcal{V}}|_{H^1(\Omega)}^2 + \|\mathbf{\Lambda}_{\mathcal{H}}(\boldsymbol{\xi}_{\mathcal{V}})\boldsymbol{\varphi}_{\mathcal{H}}\|_{L^2(\Omega)}^2 \\ &\le |\boldsymbol{\varphi}_{\mathcal{V}}|_{H^1(\Omega)}^2 + c_{\mathrm{norm}}^{-1} |||\boldsymbol{\mathcal{I}}_{\mathcal{H}}((1-\varkappa\boldsymbol{\xi}_{\mathcal{V}})\boldsymbol{\varphi}_{\mathcal{H}})|||^2.\end{aligned}$$

Furthermore, as we have due to $\boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$ and $\boldsymbol{R} < \frac{1}{|\varkappa|}$ the estimate

$$\begin{aligned}|||\boldsymbol{\mathcal{I}}_{\mathcal{H}}((1-\varkappa\boldsymbol{\xi}_{\mathcal{V}})\boldsymbol{\varphi}_{\mathcal{H}})|||^2 &= h\sum_{m=0}^{M} |(1-\varkappa\boldsymbol{\xi}_{\mathcal{V}}(\underline{x}_m))\boldsymbol{\varphi}_{\mathcal{H}}(\underline{x}_m)|^2 \\ &\le h\|1-\varkappa\boldsymbol{\xi}_{\mathcal{V}}\|_{L^\infty(\Omega)}^2 \sum_{m=0}^{M} |\boldsymbol{\varphi}_{\mathcal{H}}(\underline{x}_m)|^2 \\ &\le \left(1+\varkappa\|\boldsymbol{\xi}_{\mathcal{V}}\|_{L^\infty(\Omega)}\right)^2 |||\boldsymbol{\varphi}_{\mathcal{H}}|||^2 \\ &< 4|||\boldsymbol{\varphi}_{\mathcal{H}}|||^2,\end{aligned}$$

we get from (5.15)

$$\|\mathbf{\Lambda}(\boldsymbol{\xi})\boldsymbol{\varphi}\|_{\boldsymbol{\mathcal{X}}}^2 \le \max\left\{1, 4\frac{C_{\mathrm{norm}}}{c_{\mathrm{norm}}}\right\} \|\boldsymbol{\varphi}\|_{\boldsymbol{\mathcal{X}}}^2,$$

which yields the upper bound in (4.3). Analogously, we derive

$$\|\mathbf{\Lambda}(\boldsymbol{\xi})\boldsymbol{\varphi}\|_{\boldsymbol{\mathcal{X}}}^2 \ge \min\left\{1, (1-|\varkappa|\boldsymbol{R})\frac{C_{\mathrm{norm}}}{c_{\mathrm{norm}}}\right\} \|\boldsymbol{\varphi}\|_{\boldsymbol{\mathcal{X}}}^2, \tag{5.17}$$

which due to $\boldsymbol{R} < \frac{1}{|\varkappa|}$ yields the lower bound in (4.3).

For the Lipschitz continuity, we have with $\boldsymbol{\zeta} = (\boldsymbol{\zeta}_{\mathcal{V}}, \boldsymbol{\zeta}_{\mathcal{H}}) \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$, (5.9), and (5.15)

$$\begin{aligned}\|(\mathbf{\Lambda}(\boldsymbol{\xi}) - \mathbf{\Lambda}(\boldsymbol{\zeta}))\boldsymbol{\varphi}\|_{\boldsymbol{\mathcal{X}}}^2 &\le c_{\mathrm{norm}}^{-1} |||\boldsymbol{\mathcal{I}}_{\mathcal{H}}(\varkappa(\boldsymbol{\xi}_{\mathcal{V}} - \boldsymbol{\zeta}_{\mathcal{V}})\boldsymbol{\varphi}_{\mathcal{H}})|||^2 \\ &\le \varkappa^2 \frac{C_{\mathrm{norm}}}{c_{\mathrm{norm}}} \|\boldsymbol{\xi}_{\mathcal{V}} - \boldsymbol{\zeta}_{\mathcal{V}}\|_{L^\infty(\Omega)}^2 \|\boldsymbol{\varphi}_{\mathcal{H}}\|_{L^2(\Omega)}^2 \\ &\le \varkappa^2 \frac{C_{\mathrm{norm}}}{c_{\mathrm{norm}}} \|\boldsymbol{\xi} - \boldsymbol{\zeta}\|_{\boldsymbol{\mathcal{Y}}}^2 \|\boldsymbol{\varphi}\|_{\boldsymbol{\mathcal{X}}}^2.\end{aligned}$$

Similarly, we get

$$\|(\mathbf{\Lambda}(\boldsymbol{\xi}) - \mathbf{\Lambda}(\boldsymbol{\zeta}))\boldsymbol{\varphi}\|_{\boldsymbol{\mathcal{X}}}^2 \le \varkappa^2 \frac{C_{\mathrm{norm}}}{c_{\mathrm{norm}}} \|\boldsymbol{\xi} - \boldsymbol{\zeta}\|_{\boldsymbol{\mathcal{X}}}^2 \|\boldsymbol{\varphi}\|_{\boldsymbol{\mathcal{Y}}}^2.$$

This proves (4.4).

($\mathbf{A}$) As a direct consequence of (5.10), we have that $\mathbf{A}$ is skew-adjoint. In particular, this yields (4.5).

($\mathbf{F}$) To prove the assumptions on $\mathbf{F}$, let again $\boldsymbol{\xi} = (\boldsymbol{\xi}_{\mathcal{V}}, \boldsymbol{\xi}_{\mathcal{H}}) \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$. We first obtain from (5.7) and (5.9)

$$\begin{aligned}\|\mathbf{F}(\boldsymbol{\xi})\|_{\boldsymbol{\mathcal{Y}}} &= \|\boldsymbol{\mathcal{I}}_{\mathcal{H}}(\varkappa\boldsymbol{\xi}_{\mathcal{H}}^2)\|_{L^\infty(\Omega)} \\ &\le C_{\boldsymbol{\mathcal{I}},\infty}|\varkappa|\boldsymbol{R}^2 \\ &< C_{\boldsymbol{\mathcal{I}},\infty}\boldsymbol{R},\end{aligned}$$

which proves (4.6). Furthermore, let $\boldsymbol{\zeta} = (\boldsymbol{\zeta}_{\mathcal{V}}, \boldsymbol{\zeta}_{\mathcal{H}}) \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$. Using the norm equivalence (5.15) together with (5.9) implies

$$\begin{aligned}\|\mathbf{F}(\boldsymbol{\xi}) - \mathbf{F}(\boldsymbol{\zeta})\|_{\boldsymbol{\mathcal{X}}}^2 &\leq c_{\text{norm}}^{-1} \|\!|\boldsymbol{\mathcal{I}}_{\mathcal{H}}(\varkappa(\boldsymbol{\xi}_{\mathcal{H}}^2 - \boldsymbol{\zeta}_{\mathcal{H}}^2))\|\!|^2 \\ &\leq c_{\text{norm}}^{-1} \varkappa^2 \|\boldsymbol{\xi}_{\mathcal{H}} + \boldsymbol{\zeta}_{\mathcal{H}}\|_{L^\infty(\Omega)}^2 \|\!|\boldsymbol{\xi}_{\mathcal{H}} - \boldsymbol{\zeta}_{\mathcal{H}}\|\!|^2 \\ &< 4 \frac{C_{\text{norm}}}{c_{\text{norm}}} \|\boldsymbol{\xi} - \boldsymbol{\zeta}\|_{\boldsymbol{\mathcal{X}}}^2.\end{aligned}$$

As this yields (4.7), this concludes the proof.

□

Moreover, we obtain from [Ern and Guermond, 2004, Prop. 1.12] the following lemma, which addresses the approximation property the interpolation operator defined in (5.6).

Lemma 5.2. *Let $0 \leq r \leq p$. Then, there exists a constant $C > 0$ such that*

$$\|(\boldsymbol{\mathcal{I}}_{\mathcal{H}} - \mathrm{Id})\varphi\|_{L^2(\Omega)} + h|(\boldsymbol{\mathcal{I}}_{\mathcal{V}} - \mathrm{Id})\varphi|_{H^1(\Omega)} \leq Ch^{r+1}|\varphi|_{H^{r+1}(\Omega)}, \qquad \varphi \in H^{r+1}(\Omega). \tag{5.18}$$

In particular, this implies for $1 \leq r \leq p$ and $\xi = (\varphi, \psi) \in H^{r+1}(\Omega) \times H^r(\Omega)$ the estimates

$$\|(\boldsymbol{\mathcal{I}} - \mathrm{Id})\xi\|_{\mathcal{X}} \leq Ch^r\big(\|\varphi\|_{H^{r+1}(\Omega)} + \|\psi\|_{H^r(\Omega)}\big) \tag{5.19}$$

and

$$\|(\boldsymbol{\mathcal{I}} - \mathrm{Id})\xi\|_{\mathcal{X}} \leq C\|\xi\|_{H^1(\Omega)\times H^1(\Omega)}. \tag{5.20}$$

Finally, in the next section we state the error estimate for the space discretization of the one-dimensional Westervelt equation.

Error estimate

Using this discrete setting, we now apply the results from Section 4.3 to prove wellposedness of the one-dimensional, discrete Westervelt equation (5.13). Based on Corollary 4.29, we further derive an error estimate.

Theorem 5.3. *Let $d = 1$ and $p \geq 2$. For $\boldsymbol{R} \in (0, \frac{1}{|\varkappa|})$ let the assumptions of Theorem 3.3 be satisfied with $R < \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}\boldsymbol{R}$. Further, let the solution u of (3.7) satisfy*

$$u \in C^2(J_T, H^p(\Omega)) \cap C^1(J_T, H^{p+1}(\Omega)) \cap C(J_T, L^\infty(R)).$$

Then, there exists $h_0 > 0$ such that for all $h < h_0$, the solution $\boldsymbol{u}$ of the discrete Westervelt equation (5.13) satisfies

$$\boldsymbol{u} \in C^2(J_T, L^2(\Omega)) \cap C^1(J_T, H_0^1(\Omega)) \cap C(J_T, B_{L^\infty(\Omega)}(\boldsymbol{R})). \tag{5.21}$$

Furthermore, we have for $t \in J_T$ the bound

$$|u(t) - \boldsymbol{u}(t)|_{H^1(\Omega)} + \|\partial_t u(t) - \boldsymbol{v}(t)\|_{L^2(\Omega)} \leq C_u(1+t)\mathrm{e}^{Ct}h^p,$$

with constants $C_u, C > 0$ independent of h, t, and T, but C_u depending on the solution u and its derivatives.

Proof. To apply Corollary 4.29, we first compute $C_{\max}(h)$, which is defined in (4.45). Using (5.11) and (5.10), we obtain

$$\begin{aligned}(\mathbf{A}\boldsymbol{\varphi} \mid \boldsymbol{\psi})_{\boldsymbol{\mathcal{X}}} &= (\varphi_{\mathcal{H}} \mid \psi_{\mathcal{V}})_{H_0^1(\Omega)} - (\varphi_{\mathcal{V}} \mid \psi_{\mathcal{H}})_{H_0^1(\Omega)} \\ &\leq |\varphi_{\mathcal{H}}|_{H^1(\Omega)}|\psi_{\mathcal{V}}|_{H^1(\Omega)} + |\varphi_{\mathcal{V}}|_{H^1(\Omega)}|\psi_{\mathcal{H}}|_{H^1(\Omega)}\end{aligned}$$

Hence, taking the supremum over all $\boldsymbol{\psi} \in \boldsymbol{\mathcal{X}}$ with $\|\boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}} = 1$, together with the second estimate from (5.3) yields the esitmate (4.29) for

$$\boldsymbol{C}_{\mathbf{A}}(h) = C_{\text{inv}} h^{-1}. \tag{5.22}$$

With the first inverse estimate from (5.3), we therefore conclude

$$C_{\max}(h) = C_{\text{inv}}^2 h^{-\frac{3}{2}}. \tag{5.23}$$

For $\mathcal{Z} = \mathcal{Z}_{\mathcal{V}} \times \mathcal{Z}_{\mathcal{H}}$ defined in (3.13), we now prove that Assumption 4.28 holds true.

(A_3) Using (5.5), (5.8), and $p > \frac{3}{2}$, we obtain for $\zeta = (\zeta_{\mathcal{V}}, \zeta_{\mathcal{H}}) \in \mathcal{Z}$

$$\begin{aligned}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\zeta\|_{\boldsymbol{\mathcal{X}}} &= |(\boldsymbol{\mathcal{I}}_{\mathcal{V}} - \boldsymbol{\mathcal{L}}_{\mathcal{V}}^*)\zeta_{\mathcal{V}}|_{H^1(\Omega)} \\ &= \sup_{|\zeta_{\mathcal{V}}|_{H^1(\Omega)}=1} ((\boldsymbol{\mathcal{I}}_{\mathcal{V}} - \text{Id})\zeta_{\mathcal{V}} \mid \boldsymbol{\zeta}_{\mathcal{V}})_{H_0^1(\Omega)} \\ &= |(\boldsymbol{\mathcal{I}}_{\mathcal{V}} - \text{Id})\zeta_{\mathcal{V}}|_{H^1(\Omega)},\end{aligned} \tag{5.24}$$

which is again covered by (5.18).

(A_1) Due to the definition (5.4) of $\boldsymbol{\mathcal{L}}$ and $p > \frac{3}{2}$, (A_1) is a direct consequence of (A_3) and (5.18).

(A_2) The choice (5.12) of the discrete initial values yields that (A_2) follows directly from (A_3).

(A_5) From the definition (4.34) of $\mathcal{R}_{\text{A}}$, we deduce for $\zeta = (\zeta_{\mathcal{V}}, \zeta_{\mathcal{H}}) \in \mathcal{Z}$

$$\|\mathcal{R}_{\text{A}}\zeta\|_{\boldsymbol{\mathcal{X}}}^2 = |(\boldsymbol{\mathcal{I}}_{\mathcal{H}} - \boldsymbol{\mathcal{L}}_{\mathcal{V}}^*)\zeta_{\mathcal{H}}|_{H^1(\Omega)}^2 + \|(\boldsymbol{\Delta}\,\boldsymbol{\mathcal{L}}_{\mathcal{V}}^* - \boldsymbol{\mathcal{L}}_{\mathcal{H}}^*\,\Delta)\zeta_{\mathcal{V}}\|_{L^2(\Omega)}^2.$$

Note that the second term vanishes, as (5.5) and (5.10) imply for $\boldsymbol{\zeta}_{\mathcal{V}} \in \boldsymbol{\mathcal{X}}_{\mathcal{V}}$

$$\begin{aligned}((\boldsymbol{\Delta}\,\boldsymbol{\mathcal{L}}_{\mathcal{V}}^* - \boldsymbol{\mathcal{L}}_{\mathcal{H}}^*\,\Delta)\zeta_{\mathcal{V}} \mid \boldsymbol{\zeta}_{\mathcal{V}})_{L^2(\Omega)} &= -\,(\boldsymbol{\mathcal{L}}_{\mathcal{V}}^*\zeta_{\mathcal{V}} \mid \boldsymbol{\zeta}_{\mathcal{V}})_{H_0^1(\Omega)} - (\Delta\,\zeta_{\mathcal{V}} \mid \boldsymbol{\zeta}_{\mathcal{V}})_{L^2(\Omega)} \\ &= -\,(\zeta_{\mathcal{V}} \mid \boldsymbol{\zeta}_{\mathcal{V}})_{H_0^1(\Omega)} + (\zeta_{\mathcal{V}} \mid \boldsymbol{\zeta}_{\mathcal{V}})_{H_0^1(\Omega)} \\ &= 0.\end{aligned}$$

Since we have $\boldsymbol{\mathcal{I}}_{\mathcal{H}} = \boldsymbol{\mathcal{I}}_{\mathcal{V}}$, (5.24) yields

$$\|\mathcal{R}_{\text{A}}\zeta\|_{\boldsymbol{\mathcal{X}}} = |(\boldsymbol{\mathcal{I}}_{\mathcal{V}} - \text{Id})\zeta_{\mathcal{H}}|_{H^1(\Omega)}. \tag{5.25}$$

Thus, the bound follows from (5.18) due to $p > \frac{3}{2}$.

(A_7) As we consider a conforming discretization here, this assumption is trivially satisfied.

Hence, Assumption 4.28 is satisfied and Corollary 4.29 yields (5.21). Furthermore, we obtain from (4.61) for $y = (u, \partial_t u)$ and $\boldsymbol{y} = (\boldsymbol{u}, \boldsymbol{v})$ the estimate

$$\begin{aligned}\|y(t) - \boldsymbol{y}(t)\|_{\mathcal{X}} \leq \|(\mathrm{Id} - \boldsymbol{\mathcal{J}})y(t)\|_{\mathcal{X}} + C(1+t)\mathrm{e}^{Ct}\Big(\|\boldsymbol{\mathcal{J}} y_0 - \boldsymbol{y}_0\|_{\mathcal{X}} + \sup_{[0,t]}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})y\|_{\mathcal{X}} \\ + \sup_{[0,t]}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\partial_t y\|_{\mathcal{X}} + \sup_{[0,t]}\|(\mathrm{Id} - \boldsymbol{\mathcal{I}})\Lambda(y)\partial_t y\|_{\mathcal{X}} + \sup_{[0,t]}\|\mathcal{R}_{\mathrm{A}} y\|_{\mathcal{X}} + \sup_{[0,t]}\|(\mathrm{Id} - \boldsymbol{\mathcal{I}})\mathrm{F}(y)\|_{\mathcal{X}}\Big).\end{aligned}$$

Moreover, (5.25) implies for $s \in [0, t]$ the bound

$$\begin{aligned}\|\mathcal{R}_{\mathrm{A}} y(s)\|_{\mathcal{X}} &= |(\boldsymbol{\mathcal{I}}_{\mathcal{V}} - \mathrm{Id})\partial_t u(s)|_{H^1(\Omega)} \\ &\leq \|(\boldsymbol{\mathcal{I}} - \mathrm{Id})\partial_t y(s)\|_{\mathcal{X}}.\end{aligned} \tag{5.26}$$

Thus, (5.12) and (5.24) imply

$$\begin{aligned}\|y(t) - \boldsymbol{y}(t)\|_{\mathcal{X}} \leq C(1+t)\mathrm{e}^{Ct}\Big(\sup_{[0,t]}\|(\boldsymbol{\mathcal{I}} - \mathrm{Id})y\|_{\mathcal{X}} + \sup_{[0,t]}\|(\boldsymbol{\mathcal{I}} - \mathrm{Id})\partial_t y\|_{\mathcal{X}} \\ + \sup_{[0,t]}\|(\boldsymbol{\mathcal{I}} - \mathrm{Id})\Lambda(y)\partial_t y\|_{\mathcal{X}} + \sup_{[0,t]}\|(\boldsymbol{\mathcal{I}} - \mathrm{Id})\mathrm{F}(y)\|_{\mathcal{X}}\Big).\end{aligned}$$

Hence, we obtain with (5.19) due to $y = (u, \partial_t u)$ and $\boldsymbol{y} = (\boldsymbol{u}, \boldsymbol{v})$

$$\begin{aligned}&|u(t) - \boldsymbol{u}(t)|_{H^1(\Omega)} + \|\partial_t u(t) - \boldsymbol{v}(t)\|_{L^2(\Omega)} \\ &\quad\leq C(1+t)\mathrm{e}^{Ct}h^p\Big(\sup_{[0,t]}\|u\|_{H^{p+1}(\Omega)} + \sup_{[0,t]}\|\partial_t u\|_{H^{p+1}(\Omega)} + \sup_{[0,t]}\|\partial_t^2 u\|_{H^p(\Omega)} \\ &\qquad + \sup_{s_1,s_2\in[0,t_n]}\|\Lambda(y(s_1))\partial_t y(s_2)\|_{H^{p+1}(\Omega)\times H^p(\Omega)} + \sup_{[0,t]}\|\mathrm{F}(y)\|_{H^{p+1}(\Omega)\times H^p(\Omega)}\Big).\end{aligned} \tag{5.27}$$

Finally, as $H^{p+1}(\Omega) \times H^p(\Omega)$ is a Banach algebra for $p > \frac{1}{2}$, we get from the definition (3.8) of the nonlinearities

$$\begin{aligned}&\sup_{s_1,s_2\in[0,t_n]}\|\Lambda(y(s_1))\partial_t y(s_2)\|_{H^{p+1}(\Omega)\times H^p(\Omega)} \\ &\quad\leq C\big(1 + |\varkappa|\sup_{[0,t]}\|u\|_{H^p(\Omega)}\big)\big(\sup_{[0,t]}\|\partial_t u\|_{H^{p+1}(\Omega)} + \sup_{[0,t]}\|\partial_t^2 u\|_{H^p(\Omega)}\big),\end{aligned} \tag{5.28a}$$

and

$$\sup_{[0,t]}\|\mathrm{F}(y)\|_{H^{p+1}(\Omega)\times H^p(\Omega)} \leq C|\varkappa|\sup_{[0,t]}\|\partial_t u\|^2_{H^p(\Omega)}, \tag{5.28b}$$

cf. [Adams and Fournier, 2003, Thm. 4.39]. This completes the proof, with C_u given explicitly by (5.27) and (5.28). □

We conclude this section with the following remarks concerning the space discretization of the strongly damped Westervelt equation.

Remark 5.4. *In [Nikolić and Wohlmuth, 2019, Thm. 6.1], the conforming space discretization of the strongly damped Westervelt equation (3.15) with piecewise linear finite elements is studied. The authors analyze linearized problems and prove the wellposedness of the nonlinear discrete problem as well as an error estimate with Banach's fixed-point theorem. Compared to Theorem 5.3, their estimate yields stronger convergence rates. However, the analysis is tailored for the presence of strong damping, i.e., the error estimates for the linearized problems deteriorate for $b \to 0$. Hence, this result does not include the undamped case considered here.*

Remark 5.5. *Following the presentation in [Hipp et al., 2019, Sect. 4] for linear second-order wave-type problems, it is possible to incorporate damping in the quasilinear case in order to consider the strongly damped Westervelt equation (3.15). Despite the fact that this approach does not take the parabolic character of the strongly damped equation into account and hence does not recover the stronger convergence rates obtained in [Nikolić and Wohlmuth, 2019], it has the advantage of being stable for* $b \to 0$.

We now turn towards the space discretization of the undamped Westervelt equation in higher dimensions.

5.1.2 Example: Westervelt equation (2D, 3D)

As discussed at the beginning of this section and in Section 3.2, the space discretization of the Westervelt equation in 2D or 3D is more involved than the one-dimensional case. In particular, for the wellposedness result in Theorem 3.3, Ω has to be a bounded domain with C^3-boundary. Thus, an exact triangulation of Ω using triangles or quadrilaterals is impossible. Instead, we suggest to use isoparametric elements to approximate the domain, which leads to a non-conforming scheme.

We require the solution of the Westervelt equation (3.7) to be continuous in space, in order to apply a nodal interpolation operator to it. Thus, we see that Theorem 3.3 is only sufficient for the case $d = 1$, due to Sobolev's embedding (2.5). To circumvent this, we use the refined space $\mathcal{Y} = \mathcal{Y}_\mathcal{V} \times \mathcal{Y}_\mathcal{H}$ given by (3.12). As discussed at the end of Section 3.2, we have that Assumption 3.1 is still satisfied. Moreover, we assume that the statement of Theorem 3.3 is also true for the refined space $\mathcal{Y}$. In particular, this includes the bound

$$\|(u(t), \partial_t u(t))\|_\mathcal{Y} < R, \qquad t \in J_T.$$

Based on this assumption, we first introduce the approximation spaces and discrete operators using isoparametric elements. As Assumption 4.1 follows as in the one-dimensional case, we only have to bound the additional error terms arising from the non-conformity of the space discretization. To conclude, we use again the abstract theory from Section 4.3 and Section 4.4 to derive an error estimate.

Approximation spaces

We now introduce the discrete spaces $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$ based on isoparametric elements. We closely follow the presentation in [Hipp, 2017, Chap. 7.1], but only review the results required for the Westervelt equation. For further insight into the discretization with isoparametric elements, we refer to [Bernardi, 1989] and [Elliott and Ranner, 2013].

Throughout Section 5.1.2, we assume for some $p \geq 3$ that $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, is a bounded domain with C^{p+1}-boundary. In the following, we sketch the construction of the underlying triangulation for the isoparametric elements. For the first step, let $\Omega^\# \subset \mathbb{R}^d$ be a polygonal approximation of Ω with corresponding simplicial triangulation $\mathcal{T}_h^\#$. More precisely, $\mathcal{T}_h^\#$ consists of closed triangles for $d = 2$ or closed tetrahedra for $d = 3$. We further assume the following properties of $\mathcal{T}_h^\#$.

- For all vertices $\widehat{x} \in \partial\Omega^{\#}$, we have $\widehat{x} \in \partial\Omega$, i.e., every vertex on the boundary of the polygonal approximation $\Omega^{\#}$ lies on the boundary of the exact domain Ω.
- For each element $K_h^{\#} \in \mathcal{T}_h^{\#}$, we have that $K_h^{\#} \cap \partial\Omega^{\#}$ contains at most one face.
- The space discretization parameter

$$h := \max\{\operatorname{diam}(K_h^{\#}) \mid K_h^{\#} \in \mathcal{T}_h^{\#}\}$$

  is chosen sufficiently small such that for every $x^{\#} \in \partial\Omega^{\#}$, there exists $x \in \partial\Omega$ uniquely defined by the normal projection from $\partial\Omega^{\#}$ to $\partial\Omega$, cf. [Elliott and Ranner, 2013, Sect. 2.1] for further details.
- $\mathcal{T}_h^{\#}$ is a quasi-uniform triangulation, i.e., there exists a constant $\rho > 0$ independent of h such that

$$\min\{\operatorname{diam}(B_{K_h^{\#}}) \mid K_h^{\#} \in \mathcal{T}_h^{\#}\} \geq \rho h$$

  holds with $B_{K_h^{\#}} \subset K_h^{\#}$ being the largest ball contained in $K_h^{\#}$.

In Figure 5.1 on the left, a section of the simplicial triangulation $\mathcal{T}_h^{\#}$ (green) is illustrated for an exemplary domain Ω (gray). Since $\mathcal{T}_h^{\#}$ contains only triangles with straight edges, the boundary of Ω (black) is approximated by a piecewise linear curve.

We further employ a triangulation $\mathcal{T}_h^{\text{ex}}$ of Ω, which is constructed based on $\mathcal{T}_h^{\#}$ as explained in [Elliott and Ranner, 2013, Sect. 4.1.1]. Since the elements $K_h^{\text{ex}} \in \mathcal{T}_h^{\text{ex}}$ satisfy

$$\bigcup_{K_h^{\text{ex}} \in \mathcal{T}_h^{\text{ex}}} K_h^{\text{ex}} = \overline{\Omega},$$

$\mathcal{T}_h^{\text{ex}}$ is called exact triangulation. Moreover, note that all interior elements of $\mathcal{T}_h^{\#}$, i.e., the elements of $\mathcal{T}_h^{\#}$ which have at most one vertex at $\partial\Omega^{\#}$, are also present in $\mathcal{T}_h^{\text{ex}}$. This is also illustrated in Figure 5.1 in the center, where a section of $\mathcal{T}_h^{\text{ex}}$ (blue) is depicted.

Furthermore, if $\widehat{K} \subset \mathbb{R}^d$ denotes the reference simplex, then for every element $K_h^{\text{ex}} \in \mathcal{T}_h^{\text{ex}}$ there exists a smooth transformation

$$F_{K_h^{\text{ex}}} : \widehat{K} \to \mathbb{R}^d, \qquad K_h^{\text{ex}} = F_{K_h^{\text{ex}}}(\widehat{K}).$$

To construct the computational domain Ω_h, let $\widehat{M} \in \mathbb{N}$ and $\{\widehat{\phi}_0, \dots, \widehat{\phi}_{\widehat{M}}\}$ be the local basis of the polynomial space $\mathcal{P}^p(\widehat{K})$ on $\widehat{K}$ of degree p, which consists of Lagrange polynomials corresponding to the Gauss–Lobatto quadrature points $\{\widehat{\underline{x}}_0, \dots, \widehat{\underline{x}}_{\widehat{M}}\}$. We then define for $K_h^{\text{ex}} \in \mathcal{T}_h^{\text{ex}}$ the polynomial interpolation $F_{K_h^{\text{ex}}}^{\text{ip}}$ of $F_{K_h^{\text{ex}}}$ of degree p by

$$F_{K_h^{\text{ex}}}^{\text{ip}}(x) := \sum_{m=0}^{\widehat{M}} F_{K_h^{\text{ex}}}(\widehat{\underline{x}}_m)\widehat{\phi}_m, \qquad x \in \widehat{K}.$$

Further, we set $K_h := F_{K_h^{\text{ex}}}^{\text{ip}}(\widehat{K}) \approx K_h^{\text{ex}}$ and introduce the simplified notation $F_{K_h}(\widehat{K}) := K_h$. Finally, we define the triangulation of isoparametric elements of degree p

$$\mathcal{T}_h := \{K_h = F_{K_h^{\text{ex}}}^{\text{ip}}(\widehat{K}) \mid K_h^{\text{ex}} \in \mathcal{T}_h^{\text{ex}}\} \tag{5.29}$$

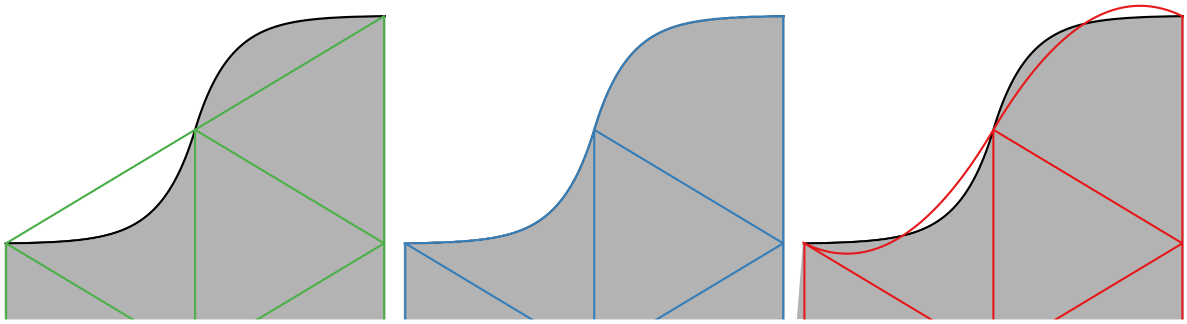

Figure 5.1: Illustration of the simplicial triangulation $\mathcal{T}_h^{\#}$ (left, green), the exact triangulation $\mathcal{T}_h^{\text{ex}}$ (center, blue), and the triangulation of quadratic isoparametric elements $\mathcal{T}_h$ (right, red).

and the computational domain

$$\Omega_h := \operatorname{int}\Big(\bigcup_{K_h \in \mathcal{T}_h} K_h\Big), \tag{5.30}$$

where int denotes the interior. By construction, $\mathcal{T}_h$ is a quasi-uniform mesh of Ω_h and $\Omega_h \approx \Omega$.

In Figure 5.1 on the right, a section of the triangulation $\mathcal{T}_h$ of isoparametric elements of degree 2 (red) is depicted. In particular, we observe that the boundary of Ω (black) is approximated by a piecewise quadratic curve.

The definitions (5.29) and (5.30) allow us to introduce the discrete approximation spaces

$$\begin{aligned} \boldsymbol{\mathcal{V}}_{\mathcal{H}} &:= \{\varphi \in C(\Omega_h) \mid \varphi|_{K_h} = \widehat{\varphi} \circ (F_{K_h})^{-1} \text{ with } \widehat{\varphi} \in \mathcal{P}^p(\widehat{K}) \text{ for } K_h \in \mathcal{T}_h\} \subset H^1(\Omega_h), \\ \boldsymbol{\mathcal{V}}_{\mathcal{V}} &:= \boldsymbol{\mathcal{V}}_{\mathcal{H}} \cap H_0^1(\Omega_h). \end{aligned} \tag{5.31}$$

As in the one-dimensional case in Section 5.1.1, based on the local bases we obtain for some $M \in \mathbb{N}$ the global basis $\{\boldsymbol{\phi}_0, \dots, \boldsymbol{\phi}_M\}$ of $\boldsymbol{\mathcal{V}}_{\mathcal{H}}$ with corresponding nodes $\{\underline{x}_0, \dots, \underline{x}_M\}$. These basis functions are later used for the construction of the interpolation operator and the discrete norm.

Moreover, for $\boldsymbol{\mathcal{X}}_{\mathcal{V}} = \boldsymbol{\mathcal{V}}_{\mathcal{V}} = \boldsymbol{\mathcal{Y}}_{\mathcal{V}}$ and $\boldsymbol{\mathcal{X}}_{\mathcal{H}} = \boldsymbol{\mathcal{V}}_{\mathcal{H}} = \boldsymbol{\mathcal{Y}}_{\mathcal{H}}$ we set $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{X}}_{\mathcal{V}} \times \boldsymbol{\mathcal{X}}_{\mathcal{H}}$ and $\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{Y}}_{\mathcal{V}} \times \boldsymbol{\mathcal{Y}}_{\mathcal{H}}$. However, in contrary to the one-dimensional case, we now have in general $\boldsymbol{\mathcal{V}} \not\subset \mathcal{X}$. We define the inner product on $\boldsymbol{\mathcal{X}}$ by

$$(\boldsymbol{\varphi} \mid \boldsymbol{\psi})_{\boldsymbol{\mathcal{X}}} = (\varphi_{\mathcal{V}} \mid \psi_{\mathcal{V}})_{H_0^1(\Omega_h)} + (\varphi_{\mathcal{H}} \mid \psi_{\mathcal{H}})_{L^2(\Omega_h)}, \qquad \boldsymbol{\varphi} = \begin{pmatrix} \varphi_{\mathcal{V}} \\ \varphi_{\mathcal{H}} \end{pmatrix}, \boldsymbol{\psi} = \begin{pmatrix} \psi_{\mathcal{V}} \\ \psi_{\mathcal{H}} \end{pmatrix} \in \boldsymbol{\mathcal{V}},$$

and the norm in $\boldsymbol{\mathcal{Y}}$ by

$$\|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{Y}}}^2 = \|\xi_{\mathcal{V}}\|_{L^\infty(\Omega_h)}^2 + \|\xi_{\mathcal{H}}\|_{L^\infty(\Omega_h)}^2, \qquad \boldsymbol{\xi} = \begin{pmatrix} \xi_{\mathcal{V}} \\ \xi_{\mathcal{H}} \end{pmatrix} \in \boldsymbol{\mathcal{V}},$$

i.e., as in (5.1) and (5.2), with Ω replaced by Ω_h.

For the inverse estimates (4.1), we provide global estimates based on the local inverse estimate from [Brenner and Scott, 2008, Lem. 4.5.3].

Lemma 5.6. *Let $d \in \{2,3\}$ and $\mathcal{T}_h$ as in (5.29), with reference element $\widehat{K} \in \mathbb{R}^d$. There exist constants $C_{\mathrm{inv},1}, C_{\mathrm{inv},2} > 0$, which may depend on $\widehat{K}$ and $\mathcal{P}^p(\widehat{K})$, but not on h such that for $\boldsymbol{\varphi} \in \boldsymbol{\mathcal{V}}_{\mathcal{H}}$, we have*

$$\|\boldsymbol{\varphi}\|_{L^\infty(\Omega_h)} \leq C_{\mathrm{inv},1} h^{-\frac{d}{2}} \|\boldsymbol{\varphi}\|_{L^2(\Omega_h)}, \qquad |\boldsymbol{\varphi}|_{H^1(\Omega_h)} \leq C_{\mathrm{inv},2} h^{-1} \|\boldsymbol{\varphi}\|_{L^2(\Omega_h)}. \tag{5.32}$$

Proof. Let $\boldsymbol{\varphi} \in \boldsymbol{\mathcal{V}}_{\mathcal{H}}$. Since $\mathcal{T}_h$ is a quasi-uniform triangulation, [Brenner and Scott, 2008, Lem. 4.5.3] yields the existence of $C_{\mathrm{inv},1}$ with

$$\|\boldsymbol{\varphi}\|_{L^\infty(K_h)} \leq C_{\mathrm{inv},1} h^{-\frac{d}{2}} \|\boldsymbol{\varphi}\|_{L^2(K_h)}, \qquad K_h \in \mathcal{T}_h.$$

As there exists $K_h^* \in \mathcal{T}_h$ with

$$\|\boldsymbol{\varphi}\|_{L^\infty(K_h^*)} = \|\boldsymbol{\varphi}\|_{L^\infty(\Omega_h)},$$

we finally derive the global estimate

$$\|\boldsymbol{\varphi}\|_{L^\infty(\Omega_h)} = \|\boldsymbol{\varphi}\|_{L^\infty(K_h^*)} \leq C_{\mathrm{inv},1} h^{-\frac{d}{2}} \|\boldsymbol{\varphi}\|_{L^2(K_h^*)} \leq C_{\mathrm{inv},1} h^{-\frac{d}{2}} \|\boldsymbol{\varphi}\|_{L^2(\Omega_h)},$$

which corresponds to the first estimate in (5.32).

For the second estimate, we again obtain from [Brenner and Scott, 2008, Lem. 4.5.3] the existence of $C_{\mathrm{inv},2}$ with

$$|\boldsymbol{\varphi}|_{H^1(K_h)} \leq C_{\mathrm{inv},2} h^{-1} \|\boldsymbol{\varphi}\|_{L^2(K_h)}, \qquad K_h \in \mathcal{T}_h.$$

This yields the global estimate

$$\Big(\sum_{K_h \in \mathcal{T}_h} |\boldsymbol{\varphi}|^2_{H^1(K_h)} \Big)^{\frac{1}{2}} \leq C_{\mathrm{inv},2} h^{-1} \Big(\sum_{K_h \in \mathcal{T}_h} \|\boldsymbol{\varphi}\|^2_{L^2(K_h)} \Big)^{\frac{1}{2}},$$

which completes the proof. □

Thus we have (4.1) with constants $\boldsymbol{C}_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}}(h) = C(h^{-1}+1)$ and $\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h) = Ch^{-\frac{d}{2}}$, for some constant $C > 0$ independent of h.

Discrete operators

We now focus on the discrete operators. As the discretization is non-conforming, the introduction of a non-trivial lift operator is essential. We follow the approach presented in [Elliott and Ranner, 2013, Sect. 4.2].

Based on the transformations $F_{K_h^{\mathrm{ex}}}$ and F_{K_h}, we obtain the elementwise smooth diffeomorphism

$$\mathcal{G} : \Omega_h \to \Omega, \qquad \mathcal{G}|_{K_h} := F_{K_h^{\mathrm{ex}}} \circ (F_{K_h})^{-1}, \qquad K_h \in \mathcal{T}_h.$$

As shown in [Elliott and Ranner, 2013, Lem. 4.6], we have

$$\mathcal{G}|_{K_h} \in C^{p+1}(K_h), \qquad K_h \in \mathcal{T}_h.$$

Thus, we define the lift operator $\boldsymbol{\mathcal{L}} : \boldsymbol{\mathcal{X}} \to \mathcal{X}$ by

$$\boldsymbol{\mathcal{L}} = \begin{pmatrix} \boldsymbol{\mathcal{L}}_{\mathcal{V}} & 0 \\ 0 & \boldsymbol{\mathcal{L}}_{\mathcal{H}} \end{pmatrix}, \tag{5.33}$$

with

$$\boldsymbol{\mathcal{L}}_{\mathcal{V}} \boldsymbol{\varphi} = \boldsymbol{\varphi} \circ \mathcal{G}^{-1}, \qquad \boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}}, \tag{5.34}$$

and $\boldsymbol{\mathcal{L}}_{\mathcal{H}} = \boldsymbol{\mathcal{L}}_{\mathcal{V}}$. From [Elliott and Ranner, 2013, Prop. 4.9], we get the existence of constants $c_{\boldsymbol{\mathcal{L}}}^*, C_{\boldsymbol{\mathcal{L}}}^* > 0$ with

$$\begin{aligned} c_{\boldsymbol{\mathcal{L}}}^* |\boldsymbol{\mathcal{L}}_{\mathcal{V}} \boldsymbol{\varphi}_{\mathcal{V}}|_{H^1(\Omega)} &\leq |\boldsymbol{\varphi}_{\mathcal{V}}|_{H^1(\Omega_h)} \leq C_{\boldsymbol{\mathcal{L}}_{\mathcal{V}}}^* |\boldsymbol{\mathcal{L}}_{\mathcal{V}} \boldsymbol{\varphi}_{\mathcal{V}}|_{H^1(\Omega)}, \\ c_{\boldsymbol{\mathcal{L}}}^* \|\boldsymbol{\mathcal{L}}_{\mathcal{H}} \boldsymbol{\varphi}_{\mathcal{H}}\|_{L^2(\Omega)} &\leq \|\boldsymbol{\varphi}_{\mathcal{H}}\|_{L^2(\Omega_h)} \leq C_{\boldsymbol{\mathcal{L}}}^* \|\boldsymbol{\mathcal{L}}_{\mathcal{H}} \boldsymbol{\varphi}_{\mathcal{H}}\|_{L^2(\Omega)}, \end{aligned} \qquad \boldsymbol{\varphi} = \begin{pmatrix} \boldsymbol{\varphi}_{\mathcal{V}} \\ \boldsymbol{\varphi}_{\mathcal{H}} \end{pmatrix} \in \boldsymbol{\mathcal{X}}. \tag{5.35}$$

This implies (4.19) with $\boldsymbol{C}_{\boldsymbol{\mathcal{L}}} = \left(c_{\boldsymbol{\mathcal{L}}}^*\right)^{-1}$. Hence, we define

$$\boldsymbol{\mathcal{L}}_{\mathcal{X}}^* = \begin{pmatrix} \boldsymbol{\mathcal{L}}_{\mathcal{V}}^* & 0 \\ 0 & \boldsymbol{\mathcal{L}}_{\mathcal{H}}^* \end{pmatrix} \tag{5.36}$$

as in (4.20).

For the interpolation operator, let $\boldsymbol{\mathcal{I}}_{\mathcal{V}} : C(\overline{\Omega}) \to C(\overline{\Omega_h})$ be the Lagrange interpolation operator with respect to the nodes $\{\underline{x}_0, \ldots, \underline{x}_M\}$. Then, we obtain as in the one-dimensional case, cf. (5.7),

$$\|\boldsymbol{\mathcal{I}}_{\mathcal{V}} \xi\|_{L^\infty(\Omega_h)} \leq C_{\boldsymbol{\mathcal{I}},\infty} \|\xi\|_{L^\infty(\Omega)}, \qquad \xi \in C(\overline{\Omega}),$$

with a constant $C_{\boldsymbol{\mathcal{I}},\infty} > 0$ independent of h. We then set for $\boldsymbol{\mathcal{I}}_{\mathcal{H}} = \boldsymbol{\mathcal{I}}_{\mathcal{V}}$

$$\boldsymbol{\mathcal{I}} = \begin{pmatrix} \boldsymbol{\mathcal{I}}_{\mathcal{V}} & 0 \\ 0 & \boldsymbol{\mathcal{I}}_{\mathcal{H}} \end{pmatrix}, \qquad \boldsymbol{\mathcal{J}} = \begin{pmatrix} \boldsymbol{\mathcal{L}}_{\mathcal{V}}^* & 0 \\ 0 & \boldsymbol{\mathcal{I}}_{\mathcal{H}} \end{pmatrix}. \tag{5.37}$$

The boundedness (4.18) of $\boldsymbol{\mathcal{I}}$ follows again from Sobolev's embedding (2.5) and the boundedness (4.17) of $\boldsymbol{\mathcal{J}}$ follows from the boundedness (4.22) of the adjoint lift $\boldsymbol{\mathcal{L}}_{\mathcal{X}}^*$.

For the discretization of the operators appearing in the Westervelt equation, we follow the approach presented in Section 4.4, i.e., we define for some $\boldsymbol{R} > 0$ specified in the next section the discrete operators as in (5.9) and (5.11), where $\boldsymbol{\Lambda}_{\mathcal{H}}(\boldsymbol{\xi}_{\mathcal{V}}) : \boldsymbol{\mathcal{X}}_{\mathcal{H}} \to \boldsymbol{\mathcal{X}}_{\mathcal{H}}$ and $\mathbf{F}_{\mathcal{H}}(\boldsymbol{\xi}_{\mathcal{H}})$ are for $\boldsymbol{\xi} = (\boldsymbol{\xi}_{\mathcal{V}}, \boldsymbol{\xi}_{\mathcal{H}}) \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$ given by

$$\boldsymbol{\Lambda}_{\mathcal{H}}(\boldsymbol{\xi}_{\mathcal{V}}) \boldsymbol{\zeta} = \boldsymbol{\mathcal{I}}_{\mathcal{H}}((1 - \varkappa \boldsymbol{\mathcal{L}} \boldsymbol{\xi}_{\mathcal{V}}) \boldsymbol{\mathcal{L}} \boldsymbol{\zeta}), \qquad \mathbf{F}_{\mathcal{H}}(\boldsymbol{\xi}_{\mathcal{H}}) = \boldsymbol{\mathcal{I}}_{\mathcal{H}}(\varkappa (\boldsymbol{\mathcal{L}} \boldsymbol{\xi}_{\mathcal{H}})^2), \qquad \boldsymbol{\zeta} \in \boldsymbol{\mathcal{X}}_{\mathcal{H}}. \tag{5.38}$$

Furthermore, the discrete Laplacian $\boldsymbol{\Delta} : \boldsymbol{\mathcal{X}}_{\mathcal{V}} \to \boldsymbol{\mathcal{X}}_{\mathcal{H}}$ is given by

$$(\boldsymbol{\Delta}\, \boldsymbol{\varphi} \mid \boldsymbol{\psi})_{L^2(\Omega_h)} := -\left(\nabla\, \boldsymbol{\varphi} \mid \nabla\, \boldsymbol{\psi}\right)_{L^2(\Omega_h)}, \qquad \boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}}_{\mathcal{V}},\ \boldsymbol{\psi} \in \boldsymbol{\mathcal{X}}_{\mathcal{H}}. \tag{5.39}$$

To conclude, we define the initial values as in (5.12). Finally, we end up with the discrete Westervelt equation in the multi-dimensional case, which corresponds to (5.13).

Properties of the discrete operators

In this section, we investigate the discrete operators defined above. More precisely, we introduce the multi-dimensional variant of the discrete norm defined in (5.14). Then, Lemma 5.1 yields again that the discrete operators satisfy Assumption 4.1 for a suitable choice of $\boldsymbol{R}$. We further review the approximation property of the interpolation operators. Finally, we derive bounds for the differences given in (4.57).

Analogous to (5.14) we define the discrete norm

$$|||\boldsymbol{\varphi}_{\mathcal{H}}|||^2_{\boldsymbol{\mathcal{X}}_{\mathcal{H}}} = h^d \sum_{m=0}^{M} |\boldsymbol{\varphi}_{\mathcal{H}}(\underline{x}_m)|^2, \qquad \boldsymbol{\varphi}_{\mathcal{H}} \in \boldsymbol{\mathcal{X}}_{\mathcal{H}}.$$

Since (5.15) and (5.16) are also satisfied in this case, we obtain that Lemma 5.1 is also satisfied for (5.38) and (5.39) instead of (5.9) and (5.10). In particular, the discrete operators defined above satisfy Assumption 4.1 for $\boldsymbol{R} \in (0, \frac{1}{|\varkappa|})$, with $\varkappa$ introduced in (3.6).

Furthermore, [Elliott and Ranner, 2013, Prop. 5.4] yields the following approximation property for the interpolation operators with isoparametric elements.

Lemma 5.7. *Let $1 \le r \le p$. Then, there exists a constant $C > 0$ such that*

$$\|(\boldsymbol{\mathcal{L}}_{\mathcal{H}}\boldsymbol{\mathcal{I}}_{\mathcal{H}} - \mathrm{Id})\varphi\|_{L^2(\Omega)} + h|(\boldsymbol{\mathcal{L}}_{\mathcal{V}}\boldsymbol{\mathcal{I}}_{\mathcal{V}} - \mathrm{Id})\varphi|_{H^1(\Omega)} \le Ch^{r+1}|\varphi|_{H^{r+1}(\Omega)}, \quad \varphi \in H^{r+1}(\Omega). \tag{5.40}$$

In particular, this implies for $2 \le r \le p$ and $\xi = (\varphi, \psi) \in H^{r+1}(\Omega) \times H^r$ the estimates

$$\|(\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}} - \mathrm{Id})\xi\|_{\mathcal{X}} \le Ch^r\big(\|\varphi\|_{H^{r+1}(\Omega)} + \|\psi\|_{H^r(\Omega)}\big) \tag{5.41}$$

and

$$\|(\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}} - \mathrm{Id})\xi\|_{\mathcal{X}} \le C\|\xi\|_{H^2(\Omega)\times H^2(\Omega)}. \tag{5.42}$$

Finally, we consider the differences including the lift, i.e., we consider for $\boldsymbol{\zeta} \in \boldsymbol{\mathcal{X}}$ the operator

$$\triangle^{\boldsymbol{\mathcal{L}}}_{\boldsymbol{\mathcal{X}}}(\boldsymbol{\zeta}) = \sup_{\|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}}=1} \Big((\boldsymbol{\zeta} \mid \boldsymbol{\xi})_{\boldsymbol{\mathcal{X}}} - (\boldsymbol{\mathcal{L}}\boldsymbol{\zeta} \mid \boldsymbol{\mathcal{L}}\boldsymbol{\xi})_{\mathcal{X}} \Big), \qquad \boldsymbol{\zeta}, \boldsymbol{\xi} \in \boldsymbol{\mathcal{X}},$$

cf. (4.57). From [Elliott and Ranner, 2013, Lem. 6.2], we obtain the following bound.

Lemma 5.8. *Let $\boldsymbol{\zeta} = (\zeta_{\mathcal{V}}, \zeta_{\mathcal{H}}) \in \boldsymbol{\mathcal{X}}$ and $1 \le r \le p$. There exists a constant $C > 0$ independent of h such that*

$$\triangle^{\boldsymbol{\mathcal{L}}}_{\boldsymbol{\mathcal{X}}}(\boldsymbol{\zeta}) \le Ch^r\|\boldsymbol{\zeta}\|_{\boldsymbol{\mathcal{X}}}. \tag{5.43}$$

Finally, we collect all preliminaries in the following section to state the error estimate for the space discretization of the multi-dimensional Westervelt equation.

Error estimate

Based on the abstract results presented in Section 4.3, we conclude this section with the error estimate for the space discretization of the multi-dimensional Westervelt equation, including specific convergence rates with respect to the discretization parameter h.

Theorem 5.9. *For $d \in \{2,3\}$ and $p \geq 3$, let $\Omega \subset \mathbb{R}^d$ be a bounded domain with C^{p+1}-boundary. For $\boldsymbol{R} \in (0, \frac{1}{|\varkappa|})$ let the assumptions of Theorem 3.3 be satisfied with $R < \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}\boldsymbol{R}$. Further, let the solution u of (3.7) satisfy*

$$u \in C^2(J_T, H^p(\Omega)) \cap C^1(J_T, H^{p+1}(\Omega)) \cap C(J_T, L^\infty(R)).$$

Then, there exists $h_0 > 0$ such that for all $h < h_0$, the solution $\boldsymbol{u}$ of the discrete Westervelt equation (5.13) satisfies

$$\boldsymbol{u} \in C^2(J_T, L^2(\Omega_h)) \cap C^1(J_T, H_0^1(\Omega_h)) \cap C(J_T, B_{L^\infty(\Omega_h)}(\boldsymbol{R})). \tag{5.44}$$

Furthermore, for $t \in J_T$ we have

$$|u(t) - \boldsymbol{\mathcal{L}}_{\mathcal{V}}\boldsymbol{u}(t)|_{H^1(\Omega)} + \|\partial_t u(t) - \boldsymbol{\mathcal{L}}_{\mathcal{H}}\boldsymbol{v}(t)\|_{L^2(\Omega)} \leq C_u(1+t)\mathrm{e}^{Ct}h^p, \tag{5.45}$$

with constants $C_u, C > 0$ independent of h, t, and T, but C_u depending on the solution u and its derivatives.

Proof. We closely follow the approach in the proof of Theorem 5.3, highlighting the main differences resulting from the non-conformity of the discretization.

To determine $C_{\max}(h)$, we first get from (5.11) and (5.39)

$$\begin{aligned}(\mathbf{A}\boldsymbol{\varphi} \mid \boldsymbol{\psi})_{\boldsymbol{\mathcal{X}}} &= (\boldsymbol{\varphi}_{\mathcal{H}} \mid \boldsymbol{\psi}_{\mathcal{V}})_{H_0^1(\Omega_h)} - (\boldsymbol{\varphi}_{\mathcal{V}} \mid \boldsymbol{\psi}_{\mathcal{H}})_{H_0^1(\Omega_h)} \\ &\leq |\boldsymbol{\varphi}_{\mathcal{H}}|_{H^1(\Omega_h)}|\boldsymbol{\psi}_{\mathcal{V}}|_{H^1(\Omega_h)} + |\boldsymbol{\varphi}_{\mathcal{V}}|_{H^1(\Omega_h)}|\boldsymbol{\psi}_{\mathcal{H}}|_{H^1(\Omega_h)}.\end{aligned}$$

Hence, taking the supremum over all $\boldsymbol{\psi} \in \boldsymbol{\mathcal{X}}$ with $\|\boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}} = 1$, together with the second estimate from (5.32) yields as in the one-dimensional case for (4.29)

$$\boldsymbol{C}_{\mathbf{A}}(h) = C_{\mathrm{inv},2}h^{-1}.$$

With the inverse estimate from (5.32), we hence conclude for (4.45)

$$C_{\max}(h) = C_{\mathrm{inv},1}C_{\mathrm{inv},2}h^{-\frac{d+2}{2}}. \tag{5.46}$$

For $\mathcal{Z} = \mathcal{Z}_{\mathcal{V}} \times \mathcal{Z}_{\mathcal{H}}$ defined in (3.13), we now check Assumption 4.28.

(A_3) From (5.37), (5.36), and (4.20), we obtain for $\zeta = (\zeta_{\mathcal{V}}, \zeta_{\mathcal{H}}) \in \mathcal{Z}$

$$\begin{aligned}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\zeta\|_{\boldsymbol{\mathcal{X}}} &= |(\boldsymbol{\mathcal{I}}_{\mathcal{V}} - \boldsymbol{\mathcal{L}}_{\mathcal{V}}^*)\zeta_{\mathcal{V}}|_{H^1(\Omega_h)} \\ &= \sup_{|\boldsymbol{\zeta}_{\mathcal{V}}|_{H^1(\Omega_h)}=1}\Big((\boldsymbol{\mathcal{I}}_{\mathcal{V}}\zeta_{\mathcal{V}} \mid \boldsymbol{\zeta}_{\mathcal{V}})_{H_0^1(\Omega_h)} - (\zeta_{\mathcal{V}} \mid \boldsymbol{\mathcal{L}}_{\mathcal{V}}\boldsymbol{\zeta}_{\mathcal{V}})_{H_0^1(\Omega)}\Big).\end{aligned}$$

With (4.57) and (4.19), this further implies

$$\begin{aligned}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\zeta\|_{\boldsymbol{\mathcal{X}}} &\leq \sup_{|\boldsymbol{\zeta}_{\mathcal{V}}|_{H^1(\Omega_h)}=1}\Big((\boldsymbol{\mathcal{I}}_{\mathcal{V}}\zeta_{\mathcal{V}} \mid \boldsymbol{\zeta}_{\mathcal{V}})_{H_0^1(\Omega_h)} - (\boldsymbol{\mathcal{L}}_{\mathcal{V}}\boldsymbol{\mathcal{I}}_{\mathcal{V}}\zeta_{\mathcal{V}} \mid \boldsymbol{\mathcal{L}}_{\mathcal{V}}\boldsymbol{\zeta}_{\mathcal{V}})_{H_0^1(\Omega)}\Big) \\ &\quad + \sup_{|\boldsymbol{\zeta}_{\mathcal{V}}|_{H^1(\Omega_h)}=1}((\boldsymbol{\mathcal{L}}_{\mathcal{V}}\boldsymbol{\mathcal{I}}_{\mathcal{V}} - \mathrm{Id})\zeta_{\mathcal{V}} \mid \boldsymbol{\mathcal{L}}_{\mathcal{V}}\boldsymbol{\zeta}_{\mathcal{V}})_{H_0^1(\Omega)} \\ &\leq \triangle^{\boldsymbol{\mathcal{L}}}_{H^1(\Omega_h)}(\boldsymbol{\mathcal{I}}_{\mathcal{V}}\zeta_{\mathcal{V}}) + \boldsymbol{C}_{\boldsymbol{\mathcal{L}}}|(\mathrm{Id} - \boldsymbol{\mathcal{L}}_{\mathcal{V}}\boldsymbol{\mathcal{I}}_{\mathcal{V}})\zeta_{\mathcal{V}}|_{H^1(\Omega)}.\end{aligned} \tag{5.47}$$

Thus, since we have $p > \frac{d+2}{2}$, we get that (5.40) and (5.43) yield (A_3).

(A_1) Due to the boundedness (4.19) of $\mathcal{L}$, we obtain (A_1) as a direct consequence of (A_3) and (5.40) since $p > \frac{d+2}{2}$.

(A_2) As the discrete initial values are defined as in the one-dimensional case, cf. (5.12), (A_2) again follows directly from (A_3).

(A_5) The definition (4.34) of $\mathcal{R}_{\mathrm{A}}$ together with (5.37) and (5.36) implies for $\zeta = (\zeta_{\mathcal{V}}, \zeta_{\mathcal{H}}) \in \mathcal{Z}$

$$\|\mathcal{R}_{\mathrm{A}}\zeta\|_{\boldsymbol{\mathcal{X}}}^2 = |(\boldsymbol{\mathcal{I}}_{\mathcal{H}} - \boldsymbol{\mathcal{L}}_{\mathcal{V}}^*)\zeta_{\mathcal{H}}|_{H^1(\Omega_h)}^2 + \|(\boldsymbol{\Delta}\, \boldsymbol{\mathcal{L}}_{\mathcal{V}}^* - \boldsymbol{\mathcal{L}}_{\mathcal{H}}^*\, \Delta)\zeta_{\mathcal{V}}\|_{L^2(\Omega_h)}.$$

Note that the second term vanishes, as (5.36), (5.39), and $\boldsymbol{\mathcal{L}}_{\mathcal{V}} = \boldsymbol{\mathcal{L}}_{\mathcal{H}}$ imply for $\boldsymbol{\zeta}_{\mathcal{V}} \in \boldsymbol{\mathcal{X}}_{\mathcal{V}}$

$$\begin{aligned}((\boldsymbol{\Delta}\, \boldsymbol{\mathcal{L}}_{\mathcal{V}}^* - \boldsymbol{\mathcal{L}}_{\mathcal{H}}^*\, \Delta)\zeta_{\mathcal{V}} \mid \boldsymbol{\zeta}_{\mathcal{V}})_{L^2(\Omega_h)} &= -\,(\boldsymbol{\mathcal{L}}_{\mathcal{V}}^*\zeta_{\mathcal{V}} \mid \boldsymbol{\zeta}_{\mathcal{V}})_{H_0^1(\Omega_h)} - (\Delta\, \zeta_{\mathcal{V}} \mid \boldsymbol{\mathcal{L}}_{\mathcal{H}}\boldsymbol{\zeta}_{\mathcal{V}})_{L^2(\Omega)} \\ &= -\,(\zeta_{\mathcal{V}} \mid \boldsymbol{\mathcal{L}}_{\mathcal{V}}\boldsymbol{\zeta}_{\mathcal{V}})_{H_0^1(\Omega)} + (\zeta_{\mathcal{V}} \mid \boldsymbol{\mathcal{L}}_{\mathcal{H}}\boldsymbol{\zeta}_{\mathcal{V}})_{H_0^1(\Omega)} \\ &= 0.\end{aligned}$$

Since we have $\boldsymbol{\mathcal{I}}_{\mathcal{H}} = \boldsymbol{\mathcal{I}}_{\mathcal{V}}$, (5.47) implies

$$\|\mathcal{R}_{\mathrm{A}}\zeta\|_{\boldsymbol{\mathcal{X}}} \le \triangle_{H^1(\Omega_h)}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}_{\mathcal{V}}\zeta_{\mathcal{H}}) + \boldsymbol{C}_{\boldsymbol{\mathcal{L}}}|(\mathrm{Id} - \boldsymbol{\mathcal{L}}_{\mathcal{V}}\boldsymbol{\mathcal{I}}_{\mathcal{V}})\zeta_{\mathcal{H}}|_{H^1(\Omega)}. \tag{5.48}$$

Thus, the bound follows from (5.40) due to $p > \frac{d+2}{2}$.

(A_7) As we have $p > \frac{d+2}{2}$, (A_7) is implied by (5.43).

Thus, Assumption 4.28 is satisfied and hence Corollary 4.29 yields (5.44). Analogous to (5.26), we obtain from (5.48) together with $y = (u, \partial_t u)$ the estimate

$$\|\mathcal{R}_{\mathrm{A}}y\|_{\boldsymbol{\mathcal{X}}} \le \triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}\partial_t y) + \boldsymbol{C}_{\boldsymbol{\mathcal{L}}}\|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}})\partial_t y\|_{\mathcal{X}}.$$

Hence, for $\boldsymbol{y} = (\boldsymbol{u}, \boldsymbol{v})$ we get from (4.61) due to (4.19), (5.47), and (5.12) the estimate

$$\begin{aligned}\|y(t) - \boldsymbol{\mathcal{L}}\boldsymbol{y}(t)\|_{\mathcal{X}} \le C(1+t)\mathrm{e}^{Ct}\Big(&\sup_{[0,t]}\|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}})y\|_{\mathcal{X}} + \sup_{[0,t]}\triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}y) \\ &+ \sup_{[0,t]}\|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}})\partial_t y\|_{\mathcal{X}} + \sup_{[0,t]}\triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}\partial_t y) \\ &+ \sup_{[0,t]}\|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}})\Lambda(y)\partial_t y\|_{\mathcal{X}} + \sup_{[0,t]}\triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}\Lambda(y)\partial_t y) \\ &+ \sup_{[0,t]}\|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}})\mathrm{F}(y)\|_{\mathcal{X}} + \sup_{[0,t]}\triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}\mathrm{F}(y))\Big).\end{aligned} \tag{5.49}$$

Furthermore, we obtain for the differences including the lift with (5.43), the bound (5.35) for the lift operator, and the interpolation property (5.41) for $\xi \in \mathcal{Z}$ the estimate

$$\begin{aligned}\triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}\xi) &\le Ch^p\|\boldsymbol{\mathcal{I}}\xi\|_{\boldsymbol{\mathcal{X}}} \\ &\le Ch^p(\|(\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}} - \mathrm{Id})\xi\|_{\mathcal{X}} + \|\xi\|_{\mathcal{X}}) \\ &\le Ch^p(Ch^2\|\xi\|_{H^3(\Omega)\times H^2(\Omega)} + \|\xi\|_{\mathcal{X}}) \\ &\le Ch^p\|\xi\|_{H^3(\Omega)\times H^2(\Omega)}.\end{aligned}$$

Using this result together with (5.41) in (5.49) yields due to $p \geq 3$

$$
\begin{aligned}
&|u(t) - \boldsymbol{\mathcal{L}}_{\mathcal{V}}\boldsymbol{u}(t)|_{H^1(\Omega)} + \|\partial_t u(t) - \boldsymbol{\mathcal{L}}_{\mathcal{H}}\boldsymbol{v}(t)\|_{L^2(\Omega)} \\
&\quad \leq C(1+t)\mathrm{e}^{Ct} h^p \Big(\sup_{[0,t]} \|u\|_{H^{p+1}(\Omega)} + \sup_{[0,t]} \|\partial_t u\|_{H^{p+1}(\Omega)} + \sup_{[0,t]} \|\partial_t^2 u\|_{H^{p}(\Omega)} \\
&\qquad + \sup_{[0,t]} \|\Lambda(y)\partial_t y\|_{H^{p+1}(\Omega)\times H^p(\Omega)} + \sup_{[0,t]} \|\mathrm{F}(y)\|_{H^{p+1}(\Omega)\times H^p(\Omega)} \Big).
\end{aligned}
$$

Finally, the constant C_u and hence also the result follow as in the one-dimensional case (5.28) from the definition (3.8) of the nonlinearities, as $H^{p+1}(\Omega) \times H^p(\Omega)$ is again a Banach algebra due to $p \geq 3$. □

In comparison with the result in Theorem 5.3 for the one-dimensional case, we observe that the lower bound for the polynomial degree is increased by one for the multi-dimensional case. This is due to the dependence of Sobolev's embedding (2.5) on the spatial dimension.

Remark 5.4 and Remark 5.5 remain valid in the multi-dimensional case. As in the one-dimensional case, the conforming space discretization with piecewise linear finite elements is considered in [Nikolić and Wohlmuth, 2019], whereas an abstract approach including also non-conforming discretizations is investigated in [Hipp et al., 2019] for linear wave-type problems.

5.2 Example: Maxwell equations

In this section we present a non-conforming space discretization for the Maxwell equations with Kerr nonlinearity, which were introduced in Section 3.3. Here, the main difference with respect to the previous examples is the usage of discontinuous approximation spaces. Thus, we briefly review in the following the essential tools for the discontinuous Galerkin method. Finally, we conclude this section with an error estimate.

Approximation spaces

For the spatial discretization of the Maxwell equations, we use the same isoparametric triangulation as for the multi-dimensional Westervelt equation in Section 5.1.2. However, we drop the continuity in the construction of the approximation spaces and hence introduce the average and jump of discrete functions. Finally, we state the inverse estimates.

Throughout Section 5.2, let $p \in \mathbb{N}$ with $p \geq 3$. Further, let $\Omega \subset \mathbb{R}^3$ be a bounded domain with C^{p+4}-boundary. We define the triangulation $\mathcal{T}_h$ of isoparametric elements and the computational domain Ω_h as in (5.29) and (5.30), respectively. In addition, we define the following sets of faces.

- Let $\mathcal{F}_h^i$ denote the set of all interior faces of $\mathcal{T}_h$, i.e., for every $F_h \in \mathcal{F}_h^i$, there exist distinct elements $K_h, \widetilde{K}_h \in \mathcal{T}_h$ with $F_h = K_h \cap \widetilde{K}_h$. For every such interior face $F_h = K_h \cap \widetilde{K}_h$, we denote by ν_{F_h} the outer unit normal vector of K_h.

- Let $\mathcal{F}_h^b$ be the set of all boundary faces of $\mathcal{T}_h$, i.e., for every $F_h \in \mathcal{F}_h^b$, there exists an element $K_h \in \mathcal{T}_h$ with $F_h = K_h \cap \partial\Omega_h$. For every such boundary face $F_h = K_h \cap \partial\Omega_h$, we fix the face unit normal vector ν_{F_h} to coincide with the outer unit normal vector ν of Ω_h.

- Let $\mathcal{F}_h = \mathcal{F}_h^i \cup \mathcal{F}_h^b$ denote the set of all faces of $\mathcal{T}_h$.

The approximation space for the magnetic field is then given by

$$\boldsymbol{\mathcal{V}}_{\mathcal{V}} \coloneqq \{\boldsymbol{\varphi} \in L^2(\Omega_h)^3 \mid \boldsymbol{\varphi}|_{K_h} = \widehat{\boldsymbol{\varphi}} \circ (F_{K_h})^{-1} \text{ with } \widehat{\boldsymbol{\varphi}} \in \mathcal{P}^p\big(\widehat{K}\big)^3 \text{ for } K_h \in \mathcal{T}_h\}.$$

As the boundary condition (3.20) for the electric field is not exactly satisfied but only approximated with the discontinuous Galerkin discretization, we increase the polynomial degree for the approximation space for the electric field by one, i.e., we set

$$\boldsymbol{\mathcal{V}}_{\mathcal{H}} \coloneqq \{\boldsymbol{\varphi} \in L^2(\Omega_h)^3 \mid \boldsymbol{\varphi}|_{K_h} = \widehat{\boldsymbol{\varphi}} \circ (F_{K_h})^{-1} \text{ with } \widehat{\boldsymbol{\varphi}} \in \mathcal{P}^{p+1}\big(\widehat{K}\big)^3 \text{ for } K_h \in \mathcal{T}_h\}. \tag{5.50}$$

Not that in contrast to (5.31), the functions in $\boldsymbol{\mathcal{V}}_{\mathcal{V}}$ and $\boldsymbol{\mathcal{V}}_{\mathcal{H}}$ are allowed to be discontinuous. Thus, we introduce in the following the average and jump of functions in $\boldsymbol{\mathcal{V}}_{\mathcal{V}}$ and $\boldsymbol{\mathcal{V}}_{\mathcal{H}}$ at the faces, based on the previously fixed face unit normal vector.

In particular, let $F_h \in \mathcal{F}_h^i$ and $K_h, \widetilde{K}_h \in \mathcal{T}_h$ such that $F_h = K_h \cap \widetilde{K}_h$ and ν_{F_h} coincides with the outer unit normal vector of K_h.

- For $\boldsymbol{\varphi} \in \boldsymbol{\mathcal{V}}_{\mathcal{V}} \cup \boldsymbol{\mathcal{V}}_{\mathcal{H}}$, we define the jump of $\boldsymbol{\varphi}$ on F_h by

$$[\![\boldsymbol{\varphi}]\!]_{F_h} \coloneqq (\boldsymbol{\varphi}|_{\widetilde{K}_h})|_{F_h} - (\boldsymbol{\varphi}|_{K_h})|_{F_h}.$$

- For $\boldsymbol{\varphi} \in \boldsymbol{\mathcal{V}}_{\mathcal{V}} \cup \boldsymbol{\mathcal{V}}_{\mathcal{H}}$, we denote by

$$\{\!\{\boldsymbol{\varphi}\}\!\}_{F_h} \coloneqq \tfrac{1}{2}\Big((\boldsymbol{\varphi}|_{\widetilde{K}_h})|_{F_h} + (\boldsymbol{\varphi}|_{K_h})|_{F_h}\Big)$$

  the average of $\boldsymbol{\varphi}$ on F_h.

In the following, we use $\boldsymbol{\mathcal{V}} = \boldsymbol{\mathcal{V}}_{\mathcal{V}} \times \boldsymbol{\mathcal{V}}_{\mathcal{H}}$ as the function space for $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{X}}_{\mathcal{V}} \times \boldsymbol{\mathcal{X}}_{\mathcal{H}}$ and $\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{Y}}_{\mathcal{V}} \times \boldsymbol{\mathcal{Y}}_{\mathcal{H}}$. For the inner product of $\boldsymbol{\mathcal{X}}$, we set

$$(\boldsymbol{\varphi} \mid \boldsymbol{\psi})_{\boldsymbol{\mathcal{X}}} = (\boldsymbol{\varphi}_{\mathcal{V}} \mid \boldsymbol{\psi}_{\mathcal{V}})_{L^2(\Omega_h)^3} + (\boldsymbol{\varphi}_{\mathcal{H}} \mid \boldsymbol{\psi}_{\mathcal{H}})_{L^2(\Omega_h)^3}, \qquad \boldsymbol{\varphi} = \begin{pmatrix} \boldsymbol{\varphi}_{\mathcal{V}} \\ \boldsymbol{\varphi}_{\mathcal{H}} \end{pmatrix}, \boldsymbol{\psi} = \begin{pmatrix} \boldsymbol{\psi}_{\mathcal{V}} \\ \boldsymbol{\psi}_{\mathcal{H}} \end{pmatrix} \in \boldsymbol{\mathcal{V}}^2. \tag{5.51}$$

As in the one-dimensional setting in (5.2), we define the norm of $\boldsymbol{\mathcal{Y}}$ by

$$\|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{Y}}}^2 = \|\boldsymbol{\xi}_{\mathcal{V}}\|_{L^\infty(\Omega_h)^3}^2 + \|\boldsymbol{\xi}_{\mathcal{H}}\|_{L^\infty(\Omega_h)^3}^2, \qquad \boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{\xi}_{\mathcal{V}} \\ \boldsymbol{\xi}_{\mathcal{H}} \end{pmatrix} \in \boldsymbol{\mathcal{V}}^2.$$

As Lemma 5.6 is also valid for the approximation spaces considered here, we directly obtain (4.1) with constants $\boldsymbol{C}_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}}(h) = C$ and $\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h) = Ch^{-\frac{3}{2}}$, for some constant $C > 0$, which is independent of h.

Discrete operators

We now focus on the definition of the operators used in the discrete setting. For the operators introduced at the end of Section 4.1, we rely on the construction in Section 5.1.2, cf. (5.33), and (5.36). However, as we use vector-valued functions here, we employ these definitions componentwise.

For the construction of the interpolation operator $\boldsymbol{\mathcal{I}}$, we have to take the discontinuity of the discrete functions into account. The global basis $\{\boldsymbol{\phi}_0, \ldots, \boldsymbol{\phi}_M\}$ of the approximation space $\boldsymbol{\mathcal{V}}_{\mathcal{V}}$

as well as the set of the corresponding nodes $\{\underline{x}_0,\dots,\underline{x}_M\}$ are obtained as the union of the local bases $\{\boldsymbol{\phi}_0^{K_h},\dots,\boldsymbol{\phi}_{\widehat{M}}^{K_h}\}$ and their corresponding nodes $\{\underline{x}_0^{K_h},\dots,\underline{x}_{\widehat{M}}^{K_h}\}$ for all elements $K_h \in \mathcal{T}_h$, respectively. Based on these local bases, let $\boldsymbol{\mathcal{I}}_{\mathcal{V}}$ be the Lagrange interpolation operator on every element $K_h \in \mathcal{T}_h$. Moreover, we define $\boldsymbol{\mathcal{I}}$ as in (5.37) and $\boldsymbol{\mathcal{J}} = \boldsymbol{\mathcal{I}}$, where we employ these definitions component-wise.

We further define the nonlinear operator as in the approach presented in Section 4.4, i.e., we define for some $\boldsymbol{R} > 0$ specified in the next section

$$\boldsymbol{\Lambda}(\boldsymbol{\xi}) = \begin{pmatrix} 1 & 0 \\ 0 & \boldsymbol{\Lambda}_{\mathcal{H}}(\boldsymbol{\xi}_{\mathcal{V}}) \end{pmatrix}, \qquad \boldsymbol{\xi}_{\mathcal{V}} \in B_{\boldsymbol{\mathcal{Y}}_{\mathcal{V}}}(\boldsymbol{R}), \tag{5.52}$$

where $\boldsymbol{\Lambda}_{\mathcal{H}}(\boldsymbol{\xi}_{\mathcal{V}}) : \boldsymbol{\mathcal{X}}_{\mathcal{H}} \to \boldsymbol{\mathcal{X}}_{\mathcal{H}}$ is for $\boldsymbol{\xi}_{\mathcal{V}} \in B_{\boldsymbol{\mathcal{Y}}_{\mathcal{V}}}(\boldsymbol{R})$ given by

$$\boldsymbol{\Lambda}_{\mathcal{H}}(\boldsymbol{\xi}_{\mathcal{V}})\boldsymbol{\zeta}_{\mathcal{H}} = \boldsymbol{\mathcal{I}}_{\mathcal{H}}\Big((1+\chi|\boldsymbol{\mathcal{L}}\boldsymbol{\xi}_{\mathcal{V}}|^2)\boldsymbol{\mathcal{L}}\boldsymbol{\zeta}_{\mathcal{H}} + 2\chi((\boldsymbol{\mathcal{L}}\boldsymbol{\xi}_{\mathcal{V}}) \otimes (\boldsymbol{\mathcal{L}}\boldsymbol{\xi}_{\mathcal{V}}))\boldsymbol{\mathcal{L}}\boldsymbol{\zeta}_{\mathcal{H}}\Big), \qquad \boldsymbol{\zeta}_{\mathcal{H}} \in \boldsymbol{\mathcal{X}}_{\mathcal{H}}. \tag{5.53}$$

We further define the discrete operator

$$\mathbf{A} = \begin{pmatrix} 0 & \mathbf{A}_{\boldsymbol{\mathcal{V}}} \\ \mathbf{A}_{\boldsymbol{\mathcal{H}}} & 0 \end{pmatrix} \tag{5.54}$$

by

$$\begin{aligned} (\mathbf{A}_{\boldsymbol{\mathcal{V}}}\boldsymbol{\varphi}_{\mathcal{H}} \mid \boldsymbol{\psi}_{\mathcal{V}})_{\boldsymbol{\mathcal{X}}_{\mathcal{V}}} = &- \sum_{K_h\in\mathcal{T}_h} (\nabla\times\boldsymbol{\varphi}_{\mathcal{H}} \mid \boldsymbol{\psi}_{\mathcal{V}})_{L^2(K_h)^3} + \sum_{F_h\in\mathcal{F}_h^b} (\nu_{F_h}\times\boldsymbol{\varphi}_{\mathcal{H}} \mid \boldsymbol{\psi}_{\mathcal{V}})_{L^2(F_h)^3} \\ &- \sum_{F_h\in\mathcal{F}_h^i} (\nu_{F_h}\times[\![\boldsymbol{\varphi}_{\mathcal{H}}]\!]_{F_h} \mid \{\!\{\boldsymbol{\psi}_{\mathcal{V}}\}\!\}_{F_h})_{L^2(F_h)^3}, \end{aligned} \tag{5.55a}$$

$$\begin{aligned} (\mathbf{A}_{\boldsymbol{\mathcal{H}}}\boldsymbol{\varphi}_{\mathcal{V}} \mid \boldsymbol{\psi}_{\mathcal{H}})_{\boldsymbol{\mathcal{X}}_{\mathcal{H}}} = &\sum_{K_h\in\mathcal{T}_h} (\nabla\times\boldsymbol{\varphi}_{\mathcal{V}} \mid \boldsymbol{\psi}_{\mathcal{H}})_{L^2(K_h)^3} \\ &+ \sum_{F_h\in\mathcal{F}_h^i} (\nu_{F_h}\times[\![\boldsymbol{\varphi}_{\mathcal{V}}]\!]_{F_h} \mid \{\!\{\boldsymbol{\psi}_{\mathcal{H}}\}\!\}_{F_h})_{L^2(F_h)^3}, \end{aligned} \tag{5.55b}$$

which corresponds to the strong form of the central fluxes discretization.

To conclude we define the initial values by

$$\boldsymbol{\mathcal{H}}_0 := \boldsymbol{\mathcal{I}}_{\mathcal{V}}\mathcal{H}_0, \qquad \boldsymbol{\mathcal{E}}_0 := \boldsymbol{\mathcal{I}}_{\mathcal{H}}\mathcal{E}_0. \tag{5.56}$$

Finally, we end up with the discrete Maxwell equations with Kerr nonlinearity, i.e., we seek the solution $(\boldsymbol{\mathcal{H}},\boldsymbol{\mathcal{E}}) : [0,T] \to \boldsymbol{\mathcal{X}}$ of

$$\begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & \boldsymbol{\Lambda}_{\mathcal{H}}(\boldsymbol{\mathcal{E}}) \end{pmatrix} \begin{pmatrix} \partial_t\boldsymbol{\mathcal{H}} \\ \partial_t\boldsymbol{\mathcal{E}} \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{A}_{\boldsymbol{\mathcal{V}}} \\ \mathbf{A}_{\boldsymbol{\mathcal{H}}} & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\mathcal{H}} \\ \boldsymbol{\mathcal{E}} \end{pmatrix} \quad \text{on } J_T, \\ \boldsymbol{\mathcal{H}}(0) = \boldsymbol{\mathcal{H}}_0, \qquad \boldsymbol{\mathcal{E}}(0) = \boldsymbol{\mathcal{E}}_0. \end{cases} \tag{5.57}$$

In the next section, we prove that the operators introduced above satisfy Assumption 4.1.

Properties of the discrete operators

As in the previous examples, we employ a discrete norm to investigate the discrete operators, cf. (5.14), (5.15), and (5.16). Here, we define the discrete norm for vector-valued functions

$$|||\boldsymbol{\varphi}_{\mathcal{H}}|||^2_{\boldsymbol{\mathcal{X}}_{\mathcal{H}}} = h^3 \sum_{K_h \in \mathcal{T}_h} \sum_{\widehat{m}=0}^{\widehat{M}} \|\boldsymbol{\varphi}_{\mathcal{H}}|_{K_h}(\underline{x}_{\widehat{m}}^{K_h})\|_2^2 = h^3 \sum_{m=0}^{M} \|\boldsymbol{\varphi}_{\mathcal{H}}(\underline{x}_m)\|_2^2, \tag{5.58}$$

for $\boldsymbol{\varphi}_{\mathcal{H}} \in \boldsymbol{\mathcal{X}}_{\mathcal{H}}$, where $\|\cdot\|_2$ denotes the Euclidean norm on $\mathbb{R}^3$. We point out that we introduce a short notation in the second step, where we implicitly employ the fact that each global node $\underline{x}_m$ corresponds to a unique local node $\underline{x}_{\widehat{m}}^{K_h}$, for some $\widehat{m} \in \{0, \dots, \widehat{M}\}$ and $K_h \in \mathcal{T}_h$.

Nevertheless, the discrete norm defined in (5.58) is again equivalent to the norm of $\boldsymbol{\mathcal{X}}_{\mathcal{H}}$, i.e., uniformly in $h > 0$ there are constants $c_{\text{norm}}, C_{\text{norm}} > 0$ such that

$$c_{\text{norm}} \|\boldsymbol{\varphi}\|^2_{L^2(\Omega)^3} \le |||\boldsymbol{\varphi}|||^2 \le C_{\text{norm}} \|\boldsymbol{\varphi}\|^2_{L^2(\Omega)^3}, \qquad \boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}}_{\mathcal{H}}, \tag{5.59}$$

holds. Further, the nodal interpolation operator satisfies

$$|||\boldsymbol{\mathcal{I}}\varphi||| = |||\varphi|||, \qquad \varphi \in \mathcal{Y}_{\mathcal{H}}. \tag{5.60}$$

Using this norm, we show in the following lemma that the discrete operators satisfy Assumption 4.1. Obviously, the assumptions for $\mathbf{F}$ are trivially satisfied since $\mathbf{F} \equiv 0$.

Lemma 5.10. *Let $\boldsymbol{R} \in \big(0, (9\|\chi\|_{L^\infty(\Omega)})^{-\frac{1}{2}}\big)$. Then, the discrete operators $\boldsymbol{\Lambda}$ and $\mathbf{A}$ given in (5.52) and (5.54) satisfy Assumption 4.1.*

Proof. Throughout this proof, for the sake of readability, we use for $m \in \{0, \dots, M\}$ the superscript m to denote the point evaluation of a function at the interpolation point $\underline{x}_m$.

($\boldsymbol{\Lambda}$) Let $\boldsymbol{\xi} = (\boldsymbol{\xi}_{\mathcal{V}}, \boldsymbol{\xi}_{\mathcal{H}}) \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$ and $\boldsymbol{\varphi} = (\boldsymbol{\varphi}_{\mathcal{V}}, \boldsymbol{\varphi}_{\mathcal{H}}) \in \boldsymbol{\mathcal{X}}$. The definition (5.51) of the norm of $\boldsymbol{\mathcal{X}}$ yields

$$\|\boldsymbol{\Lambda}(\boldsymbol{\xi})\boldsymbol{\varphi}\|^2_{\boldsymbol{\mathcal{X}}} = \|\boldsymbol{\varphi}_{\mathcal{V}}\|^2_{L^2(\Omega)^3} + \|\boldsymbol{\Lambda}_{\mathcal{H}}(\boldsymbol{\xi}_{\mathcal{V}})\boldsymbol{\varphi}_{\mathcal{H}}\|^2_{L^2(\Omega)^3}.$$

Thus, the definitions (5.53) and (5.58) of the nonlinearity and the discrete norm, respectively, together with (5.59) and (5.60) imply

$$\|\boldsymbol{\Lambda}_{\mathcal{H}}(\boldsymbol{\xi}_{\mathcal{V}})\boldsymbol{\varphi}_{\mathcal{H}}\|^2_{L^2(\Omega)^3} \le c_{\text{norm}}^{-1} h^3 \sum_{m=0}^{M} \|(1 + \chi^m |\boldsymbol{\xi}_{\mathcal{V}}^m|^2 + 2\chi^m (\boldsymbol{\xi}_{\mathcal{V}}^m \otimes \boldsymbol{\xi}_{\mathcal{V}}^m))\boldsymbol{\varphi}_{\mathcal{H}}^m\|_2^2,$$

where $\|\cdot\|_2$ denotes the Euclidean norm on $\mathbb{R}^3$. Since (2.2) yields for $m \in \{0, \dots, M\}$

$$\|(1 + \chi^m |\boldsymbol{\xi}_{\mathcal{V}}^m|^2 + 2\chi^m (\boldsymbol{\xi}_{\mathcal{V}}^m \otimes \boldsymbol{\xi}_{\mathcal{V}}^m)\boldsymbol{\varphi}_{\mathcal{H}}^m\|_2 \le (1 + 9|\chi^m| \|\boldsymbol{\xi}_{\mathcal{V}}^m\|_\infty^2) \|\boldsymbol{\varphi}_{\mathcal{H}}^m\|_2,$$

we finally obtain due to $\boldsymbol{\varphi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$ with $\boldsymbol{R} < (9\|\chi\|_{L^\infty(\Omega)})^{-\frac{1}{2}}$ the bound

$$\|\boldsymbol{\Lambda}(\boldsymbol{\xi})\boldsymbol{\varphi}\|^2_{\boldsymbol{\mathcal{X}}} \le \max\Big\{1, 4\frac{C_{\text{norm}}}{c_{\text{norm}}}\Big\} \|\boldsymbol{\varphi}\|^2_{\boldsymbol{\mathcal{X}}},$$

which yields the upper bound in (4.3). Moreover, we have

$$(\boldsymbol{\Lambda}(\boldsymbol{\xi})\boldsymbol{\varphi} \mid \boldsymbol{\varphi})_{\boldsymbol{\mathcal{X}}} \geq \min\Big\{1, (1 - 9\|\chi\|_{L^\infty(\Omega)}\boldsymbol{R}^2)\frac{C_{\mathrm{norm}}}{c_{\mathrm{norm}}}\Big\}\|\boldsymbol{\varphi}\|_{\boldsymbol{\mathcal{X}}}^2, \tag{5.61}$$

which yields the lower bound in (4.3).

For the Lipschitz continuity, we derive with $\boldsymbol{\zeta} = (\boldsymbol{\zeta}_{\mathcal{V}}, \boldsymbol{\zeta}_{\mathcal{H}}) \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$, (5.53), and (5.59) the estimate

$$\|(\boldsymbol{\Lambda}(\boldsymbol{\xi}) - \boldsymbol{\Lambda}(\boldsymbol{\zeta}))\boldsymbol{\varphi}\|_{\boldsymbol{\mathcal{X}}}^2 \leq c_{\mathrm{norm}}^{-1} h^3 \sum_{m=0}^{M} \|\chi^m\Big((|\boldsymbol{\xi}_{\mathcal{V}}^m|^2 - |\boldsymbol{\zeta}_{\mathcal{V}}^m|^2) + 2(\boldsymbol{\xi}_{\mathcal{V}}^m \otimes \boldsymbol{\xi}_{\mathcal{V}}^m - \boldsymbol{\zeta}_{\mathcal{V}}^m \otimes \boldsymbol{\zeta}_{\mathcal{V}}^m)\Big)\boldsymbol{\varphi}_{\mathcal{H}}^m\|_2^2.$$

Further, (2.2) implies for the first term

$$\begin{aligned}\|\chi^m(|\boldsymbol{\xi}_{\mathcal{V}}^m|^2 - |\boldsymbol{\zeta}_{\mathcal{V}}^m|^2)\boldsymbol{\varphi}_{\mathcal{H}}^m\|_2 &= \|\chi^m\Big((\boldsymbol{\xi}_{\mathcal{V}}^m + \boldsymbol{\zeta}_{\mathcal{V}}^m) \cdot (\boldsymbol{\xi}_{\mathcal{V}}^m - \boldsymbol{\zeta}_{\mathcal{V}}^m)\Big)\boldsymbol{\varphi}_{\mathcal{H}}^m\|_2 \\ &\leq 3\|\chi^m\|_\infty \|\boldsymbol{\xi}_{\mathcal{V}}^m + \boldsymbol{\zeta}_{\mathcal{V}}^m\|_\infty \|\boldsymbol{\xi}_{\mathcal{V}}^m - \boldsymbol{\zeta}_{\mathcal{V}}^m\|_\infty \|\boldsymbol{\varphi}_{\mathcal{H}}^m\|_2,\end{aligned}$$

while (2.1) and (2.2) yield for the second term

$$\begin{aligned}&\|\chi^m(\boldsymbol{\xi}_{\mathcal{V}}^m \otimes \boldsymbol{\xi}_{\mathcal{V}}^m - \boldsymbol{\zeta}_{\mathcal{V}}^m \otimes \boldsymbol{\zeta}_{\mathcal{V}}^m)\boldsymbol{\varphi}_{\mathcal{H}}^m\|_2 \\ &\quad= \|\chi^m\Big(\boldsymbol{\xi}_{\mathcal{V}}^m(\boldsymbol{\xi}_{\mathcal{V}}^m \cdot \boldsymbol{\varphi}_{\mathcal{H}}^m) - \boldsymbol{\zeta}_{\mathcal{V}}^m(\boldsymbol{\zeta}_{\mathcal{V}}^m \cdot \boldsymbol{\varphi}_{\mathcal{H}}^m)\Big)\|_2 \\ &\quad= \|\chi^m\Big((\boldsymbol{\xi}_{\mathcal{V}}^m - \boldsymbol{\zeta}_{\mathcal{V}}^m)(\boldsymbol{\xi}_{\mathcal{V}}^m \cdot \boldsymbol{\varphi}_{\mathcal{H}}^m) + \boldsymbol{\zeta}_{\mathcal{V}}^m((\boldsymbol{\xi}_{\mathcal{V}}^m - \boldsymbol{\zeta}_{\mathcal{V}}^m) \cdot \boldsymbol{\varphi}_{\mathcal{H}}^m)\Big)\|_2 \\ &\quad\leq 3\|\chi^m\|_\infty\Big(\|\boldsymbol{\xi}_{\mathcal{V}}^m - \boldsymbol{\zeta}_{\mathcal{V}}^m\|_2 \|\boldsymbol{\xi}_{\mathcal{V}}^m\|_\infty + \|\boldsymbol{\zeta}_{\mathcal{V}}^m\|_2 \|\boldsymbol{\xi}_{\mathcal{V}}^m - \boldsymbol{\zeta}_{\mathcal{V}}^m\|_\infty\Big)\|\boldsymbol{\varphi}_{\mathcal{H}}^m\|_\infty.\end{aligned}$$

Since all p-norms are equivalent in $\mathbb{R}^3$, we obtain due to $\boldsymbol{\xi}, \boldsymbol{\zeta} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$, the definition (2.4) of norms of vector-valued functions, and the norm equivalence (5.59)

$$\begin{aligned}\|(\boldsymbol{\Lambda}(\boldsymbol{\xi}) - \boldsymbol{\Lambda}(\boldsymbol{\zeta}))\boldsymbol{\varphi}\|_{\boldsymbol{\mathcal{X}}}^2 &\leq C c_{\mathrm{norm}}^{-1} \boldsymbol{R}^2 h^3 \sum_{m=0}^{M}\Big(\|\chi^m\|_\infty \|\boldsymbol{\xi}_{\mathcal{V}}^m - \boldsymbol{\zeta}_{\mathcal{V}}^m\|_\infty \|\boldsymbol{\varphi}_{\mathcal{H}}^m\|_2\Big)^2 \\ &\leq C\frac{C_{\mathrm{norm}}}{c_{\mathrm{norm}}}\boldsymbol{R}^2 \|\chi\|_{L^\infty(\Omega)}^2 \|\boldsymbol{\xi}_{\mathcal{V}} - \boldsymbol{\zeta}_{\mathcal{V}}\|_{L^\infty(\Omega)^3}^2 \|\boldsymbol{\varphi}_{\mathcal{H}}\|_{L^2(\Omega)^3}^2.\end{aligned}$$

This implies (4.4a). Analogously, (4.4b) is shown.

(**A**) From (5.55), we get with integration by parts

$$(\mathbf{A}_{\boldsymbol{\mathcal{V}}}\boldsymbol{\varphi}_{\mathcal{H}} \mid \boldsymbol{\varphi}_{\mathcal{V}})_{\boldsymbol{\mathcal{X}}_{\mathcal{V}}} + (\mathbf{A}_{\boldsymbol{\mathcal{H}}}\boldsymbol{\varphi}_{\mathcal{V}} \mid \boldsymbol{\varphi}_{\mathcal{H}})_{\boldsymbol{\mathcal{X}}_{\mathcal{H}}} = 0, \qquad \boldsymbol{\varphi} = \begin{pmatrix}\boldsymbol{\varphi}_{\mathcal{V}} \\ \boldsymbol{\varphi}_{\mathcal{H}}\end{pmatrix} \in \boldsymbol{\mathcal{X}}.$$

Hence, $\mathbf{A}$ is skew-adjoint, which yields (4.5). □

Moreover, we emphasize that due to the construction of the triangulation $\mathcal{T}_h$ and the fact that we both approximation spaces $\boldsymbol{\mathcal{V}}_{\mathcal{V}}$ and $\boldsymbol{\mathcal{V}}_{\mathcal{H}}$ use a uniform polynomial degree on each element $K_h \in \mathcal{T}_h$, the nodal interpolation operator satisfies $\boldsymbol{\mathcal{I}} : C(\overline{\Omega})^6 \to C(\overline{\Omega_h})^6$. Furthermore, the following variant of Lemma 5.7 for vector-valued functions holds.

Lemma 5.11. *Let $1 \leq r \leq p$. Then, there exists a constant $C > 0$ such that*

$$\|\varphi - \boldsymbol{\mathcal{L}}_{\mathcal{H}} \boldsymbol{\mathcal{I}}_{\mathcal{H}} \varphi\|_{L^2(\Omega)^3} + h|\varphi - \boldsymbol{\mathcal{L}}_{\mathcal{V}} \boldsymbol{\mathcal{I}}_{\mathcal{V}} \varphi|_{H^1(\Omega)^3} \leq Ch^{r+1} |\varphi|_{H^{r+1}(\Omega)^3}, \qquad \varphi \in H^{r+1}(\Omega)^3. \tag{5.62}$$

Finally, we require bounds for the differences of inner products including the lift. First, note that Lemma 5.8 also holds in this setting, as it is based on elementwise estimates. Furthermore, we prove the following variant for differences including the lifts as well as the differential operators A and **A**.

Lemma 5.12. *Let $1 \leq r \leq p \leq$ and $\varphi = (\varphi_{\mathcal{V}}, \varphi_{\mathcal{H}}) \in \mathcal{Y}$ with $\varphi_{\mathcal{H}} \in H^{r+1}(\Omega)^3$. Then, there exists a constant $C > 0$ such that*

$$\sup_{\|\boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}}=1} \Big((\mathbf{A} \boldsymbol{\mathcal{I}} \varphi \mid \boldsymbol{\psi})_{\boldsymbol{\mathcal{X}}} - (\mathrm{A} \boldsymbol{\mathcal{L}} \boldsymbol{\mathcal{I}} \varphi \mid \boldsymbol{\mathcal{L}} \boldsymbol{\psi})_{\mathcal{X}} \Big) \leq Ch^r \big(\|\varphi\|_{\mathcal{Y}} + \|\varphi_{\mathcal{H}}\|_{H^{r+1}(\Omega)^3} \big).$$

Proof. Let $\varphi = (\varphi_{\mathcal{V}}, \varphi_{\mathcal{H}}) \in \mathcal{Y}$ and $\boldsymbol{\psi} = (\boldsymbol{\psi}_{\mathcal{V}}, \boldsymbol{\psi}_{\mathcal{H}}) \in \boldsymbol{\mathcal{X}}$. The definitions (3.22) and (5.54) of A and **A**, respectively, imply

$$\begin{aligned}
(\mathbf{A} \boldsymbol{\mathcal{I}} \varphi \mid \boldsymbol{\psi})_{\boldsymbol{\mathcal{X}}} - (\mathrm{A} \boldsymbol{\mathcal{L}} \boldsymbol{\mathcal{I}} \varphi \mid \boldsymbol{\mathcal{L}} \boldsymbol{\psi})_{\mathcal{X}} &= \sum_{F_h \in \mathcal{F}_h^i} (\nu_{F_h} \times [\boldsymbol{\mathcal{I}}_{\mathcal{V}} \varphi_{\mathcal{V}}]_{F_h} \mid \{\!\{\boldsymbol{\psi}_{\mathcal{H}}\}\!\}_{F_h})_{L^2(F_h)^3} \\
&+ \sum_{K_h \in \mathcal{T}_h} (\nabla\times \boldsymbol{\mathcal{I}}_{\mathcal{V}} \varphi_{\mathcal{V}} \mid \boldsymbol{\psi}_{\mathcal{H}})_{L^2(K_h)^3} - (\nabla\times \boldsymbol{\mathcal{L}}_{\mathcal{V}} \boldsymbol{\mathcal{I}}_{\mathcal{V}} \varphi_{\mathcal{V}} \mid \boldsymbol{\mathcal{L}}_{\mathcal{H}} \boldsymbol{\psi}_{\mathcal{H}})_{L^2(\Omega)^3} \\
&+ (\nabla\times \boldsymbol{\mathcal{L}}_{\mathcal{H}} \boldsymbol{\mathcal{I}}_{\mathcal{H}} \varphi_{\mathcal{H}} \mid \boldsymbol{\mathcal{L}}_{\mathcal{V}} \boldsymbol{\psi}_{\mathcal{V}})_{L^2(\Omega)^3} - \sum_{K_h \in \mathcal{T}_h} (\nabla\times \boldsymbol{\mathcal{I}}_{\mathcal{H}} \varphi_{\mathcal{H}} \mid \boldsymbol{\psi}_{\mathcal{V}})_{L^2(K_h)^3} \\
&- \sum_{F_h \in \mathcal{F}_h^i} (\nu_{F_h} \times [\boldsymbol{\mathcal{I}}_{\mathcal{H}} \varphi_{\mathcal{H}}]_{F_h} \mid \{\!\{\boldsymbol{\psi}_{\mathcal{V}}\}\!\}_{F_h})_{L^2(F_h)^3} + \sum_{F_h \in \mathcal{F}_h^b} (\nu_{F_h} \times \boldsymbol{\mathcal{I}}_{\mathcal{H}} \varphi_{\mathcal{H}} \mid \boldsymbol{\psi}_{\mathcal{V}})_{L^2(F_h)^3} .
\end{aligned} \tag{5.63}$$

We now consider these terms separately. To begin with, we obtain due to $\boldsymbol{\mathcal{I}}_{\mathcal{H}} \varphi_{\mathcal{H}}, \boldsymbol{\mathcal{I}}_{\mathcal{V}} \varphi_{\mathcal{V}} \in C(\Omega_h)$ for the contributions of the interior faces

$$\begin{aligned}
(\nu_{F_h} \times [\boldsymbol{\mathcal{I}}_{\mathcal{H}} \varphi_{\mathcal{H}}]_{F_h} \mid \{\!\{\boldsymbol{\psi}_{\mathcal{V}}\}\!\}_{F_h})_{L^2(F_h)^3} &= 0, \\
(\nu_{F_h} \times [\boldsymbol{\mathcal{I}}_{\mathcal{V}} \varphi_{\mathcal{V}}]_{F_h} \mid \{\!\{\boldsymbol{\psi}_{\mathcal{H}}\}\!\}_{F_h})_{L^2(F_h)^3} &= 0,
\end{aligned} \tag{5.64}$$

for $F_h \in \mathcal{F}_h^i$.

For the volume terms, we first get from definition (5.34) of the lift operators $\boldsymbol{\mathcal{L}}_{\mathcal{V}}$ and $\boldsymbol{\mathcal{L}}_{\mathcal{H}}$ the substitution rule for integration

$$\int_{\Omega_h} \boldsymbol{\xi} \boldsymbol{\zeta} \,\mathrm{d}x = \int_{\Omega} (\boldsymbol{\mathcal{L}}_{\mathcal{V}} \boldsymbol{\xi}) (\boldsymbol{\mathcal{L}}_{\mathcal{H}} \boldsymbol{\zeta}) |\det \mathrm{D}\, \mathcal{G}^{-1}| \,\mathrm{d}x, \qquad \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\zeta} \end{pmatrix} \in \boldsymbol{\mathcal{X}}, \tag{5.65}$$

where $\mathrm{D}\, \mathcal{G}^{-1}$ denotes the elementwise defined Jacobian of $\mathcal{G}^{-1}$. Therefore, as $\boldsymbol{\mathcal{I}}_{\mathcal{H}} \varphi_{\mathcal{H}}, \boldsymbol{\mathcal{I}}_{\mathcal{V}} \varphi_{\mathcal{V}} \in C(\Omega_h)$ are piecewise polynomials, we obtain for the second line of (5.63)

$$\begin{aligned}
&\sum_{K_h \in \mathcal{T}_h} (\nabla\times \boldsymbol{\mathcal{I}}_{\mathcal{V}} \varphi_{\mathcal{V}} \mid \boldsymbol{\psi}_{\mathcal{H}})_{L^2(K_h)^3} - (\nabla\times \boldsymbol{\mathcal{L}}_{\mathcal{V}} \boldsymbol{\mathcal{I}}_{\mathcal{V}} \varphi_{\mathcal{V}} \mid \boldsymbol{\mathcal{L}}_{\mathcal{H}} \boldsymbol{\psi}_{\mathcal{H}})_{\Omega} \\
&\quad = \int_{\Omega} (\boldsymbol{\mathcal{L}}_{\mathcal{V}}\, \nabla\times \boldsymbol{\mathcal{I}}_{\mathcal{V}} \varphi_{\mathcal{V}}) (\boldsymbol{\mathcal{L}}_{\mathcal{H}} \boldsymbol{\psi}_{\mathcal{H}}) |\det \mathrm{D}\, \mathcal{G}^{-1}| - (\nabla\times \boldsymbol{\mathcal{L}}_{\mathcal{V}} \boldsymbol{\mathcal{I}}_{\mathcal{V}} \varphi_{\mathcal{V}}) (\boldsymbol{\mathcal{L}}_{\mathcal{H}} \boldsymbol{\psi}_{\mathcal{H}}) \,\mathrm{d}x \\
&\quad = \int_{\Omega} (\boldsymbol{\mathcal{L}}_{\mathcal{V}}\, \nabla\times \boldsymbol{\mathcal{I}}_{\mathcal{V}} \varphi_{\mathcal{V}}) (\boldsymbol{\mathcal{L}}_{\mathcal{H}} \boldsymbol{\psi}_{\mathcal{H}}) (|\det \mathrm{D}\, \mathcal{G}^{-1}| - 1) \,\mathrm{d}x \\
&\qquad + \int_{\Omega} \Big(\boldsymbol{\mathcal{L}}_{\mathcal{V}}\, \nabla\times \boldsymbol{\mathcal{I}}_{\mathcal{V}} \varphi_{\mathcal{V}} - \nabla\times \boldsymbol{\mathcal{L}}_{\mathcal{V}} \boldsymbol{\mathcal{I}}_{\mathcal{V}} \varphi_{\mathcal{V}} \Big) (\boldsymbol{\mathcal{L}}_{\mathcal{H}} \boldsymbol{\psi}_{\mathcal{H}}) \,\mathrm{d}x.
\end{aligned} \tag{5.66}$$

Using the Cauchy–Schwarz inequality together with

$$\|\nabla\times \boldsymbol{\xi}\|_{L^2(K_h)^3} \leq 2\|\boldsymbol{\xi}\|_{H^1(K_h)^3}, \qquad \boldsymbol{\xi} \in H^1(K_h)^3, \tag{5.67}$$

and the boundedness (5.35) of the lift operators, we further deduce

$$\begin{aligned}
&\int_\Omega (\boldsymbol{\mathcal{L}}_\mathcal{V}\, \nabla\times \boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V})(\boldsymbol{\mathcal{L}}_\mathcal{H}\boldsymbol{\psi}_\mathcal{H})(|\det \mathrm{D}\,\mathcal{G}^{-1}| - 1)\,\mathrm{d}x \\
&\quad \leq \|\boldsymbol{\mathcal{L}}_\mathcal{V}\, \nabla\times \boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V}\|_{L^2(\Omega)^3}\|\boldsymbol{\mathcal{L}}_\mathcal{H}\boldsymbol{\psi}_\mathcal{H}\|_{L^2(\Omega)^3}\||\det \mathrm{D}\,\mathcal{G}^{-1}| - 1\|_{L^\infty(\Omega)} \\
&\quad \leq C\|\boldsymbol{\mathcal{L}}_\mathcal{V}\boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V}\|_{H^1(\Omega)^3}\|\boldsymbol{\psi}_\mathcal{H}\|_{L^2(\Omega_h)^3}\||\det \mathrm{D}\,\mathcal{G}^{-1}| - 1\|_{L^\infty(\Omega)}.
\end{aligned}$$

Since the chain rule and the implicit function theorem imply

$$(\nabla\times\, \rho)\circ \mathcal{G}^{-1} = (\mathrm{D}\,\mathcal{G}^T \circ \mathcal{G}^{-1})\,\nabla\times(\rho\circ\mathcal{G}^{-1}),$$

we have

$$\begin{aligned}
\|(\boldsymbol{\mathcal{L}}_\mathcal{V}\, \nabla\times \boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V}) - (\nabla\times \boldsymbol{\mathcal{L}}_\mathcal{V}\boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V})\|_{L^2(\Omega)^3} &= \|(\nabla\times \boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V})\circ\mathcal{G}^{-1} - \nabla\times(\boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V}\circ\mathcal{G}^{-1})\|_{L^2(\Omega)^3} \\
&\leq \|\mathrm{D}\,\mathcal{G}^T\circ\mathcal{G}^{-1} - \mathrm{Id}\|_{L^\infty(\Omega)^{3\times 3}}\|\boldsymbol{\mathcal{L}}_\mathcal{V}\boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V}\|_{H^1(\Omega)^3}.
\end{aligned}$$

Thus, we obtain

$$\begin{aligned}
&\int_\Omega \Big(\boldsymbol{\mathcal{L}}_\mathcal{V}\, \nabla\times \boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V} - \nabla\times \boldsymbol{\mathcal{L}}_\mathcal{V}\boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V}\Big)(\boldsymbol{\mathcal{L}}_\mathcal{H}\boldsymbol{\psi}_\mathcal{H})\,\mathrm{d}x \\
&\quad \leq \|(\boldsymbol{\mathcal{L}}_\mathcal{V}\, \nabla\times \boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V}) - (\nabla\times \boldsymbol{\mathcal{L}}_\mathcal{V}\boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V})\|_{L^2(\Omega)^3}\|\boldsymbol{\mathcal{L}}_\mathcal{H}\boldsymbol{\psi}_\mathcal{H}\|_{L^2(\Omega)^3} \\
&\quad \leq C\|\boldsymbol{\mathcal{L}}_\mathcal{V}\boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V}\|_{H^1(\Omega)^3}\|\boldsymbol{\psi}_\mathcal{H}\|_{L^2(\Omega_h)^3}\|\mathrm{D}\,\mathcal{G}^T\circ\mathcal{G}^{-1} - \mathrm{Id}\|_{L^\infty(\Omega)^{3\times 3}}.
\end{aligned}$$

Collecting these results in (5.66), we have shown

$$\begin{aligned}
&\sum_{K_h\in\mathcal{T}_h} (\nabla\times \boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V} \mid \boldsymbol{\psi}_\mathcal{H})_{L^2(K_h)^3} - (\nabla\times \boldsymbol{\mathcal{L}}_\mathcal{V}\boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V} \mid \boldsymbol{\mathcal{L}}_\mathcal{H}\boldsymbol{\psi}_\mathcal{H})_{L^2(\Omega)^3} \\
&\quad \leq C\|\boldsymbol{\mathcal{L}}_\mathcal{V}\boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V}\|_{H^1(\Omega)^3}\|\boldsymbol{\psi}_\mathcal{H}\|_{L^2(\Omega_h)^3}\big(\||\det \mathrm{D}\,\mathcal{G}^{-1}| - 1\|_{L^\infty(\Omega)} + \|\mathrm{D}\,\mathcal{G}^T\circ\mathcal{G}^{-1} - \mathrm{Id}\|_{L^\infty(\Omega)^{3\times 3}}\big).
\end{aligned}$$

[Hipp, 2017, Lem. 7.3] provides the estimate

$$\||\det \mathrm{D}\,\mathcal{G}^{-1}| - 1\|_{L^\infty(\Omega)} + \|\mathrm{D}\,\mathcal{G}^T\circ\mathcal{G}^{-1} - \mathrm{Id}\|_{L^\infty(\Omega)^{3\times 3}} \leq Ch^r,$$

so that we finally obtain from the interpolation estimate (5.62)

$$\sum_{K_h\in\mathcal{T}_h} (\nabla\times \boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V} \mid \boldsymbol{\psi}_\mathcal{H})_{L^2(K_h)^3} - (\nabla\times \boldsymbol{\mathcal{L}}_\mathcal{V}\boldsymbol{\mathcal{I}}_\mathcal{V}\varphi_\mathcal{V} \mid \boldsymbol{\mathcal{L}}_\mathcal{H}\boldsymbol{\psi}_\mathcal{H})_{L^2(\Omega)^3} \leq Ch^r\|\varphi\|_\mathcal{Y}\|\boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}}. \tag{5.68}$$

The third line in (5.63) can be treated equivalently.

We now consider the contributions of the boundary faces in (5.63). Due to the boundary condition (3.20) of a perfect conductor, we have $\varphi_\mathcal{H}\times\nu = 0$. Thus, the substitution rule for integration (5.65) yields

$$\begin{aligned}
\sum_{F_h\in\mathcal{F}_h^b} (\nu_{F_h}\times \boldsymbol{\mathcal{I}}_\mathcal{H}\varphi_\mathcal{H} \mid \psi_\mathcal{V})_{L^2(F_h)^3} &= \int_{\partial\Omega} (\boldsymbol{\mathcal{L}}_\mathcal{H}\nu_{\Omega_h})\times(\boldsymbol{\mathcal{L}}_\mathcal{H}\boldsymbol{\mathcal{I}}_\mathcal{H}\varphi_\mathcal{H})(\boldsymbol{\mathcal{L}}_\mathcal{V}\boldsymbol{\psi}_\mathcal{V})|\det \mathrm{D}\,\mathcal{G}^{-1}|\,\mathrm{d}x \\
&= \int_{\partial\Omega} (\boldsymbol{\mathcal{L}}_\mathcal{H}\nu_{\Omega_h})\times(\boldsymbol{\mathcal{L}}_\mathcal{H}\boldsymbol{\mathcal{I}}_\mathcal{H}\varphi_\mathcal{H})(\boldsymbol{\mathcal{L}}_\mathcal{V}\boldsymbol{\psi}_\mathcal{V})(|\det \mathrm{D}\,\mathcal{G}^{-1}| - 1)\,\mathrm{d}x \\
&\quad + \big((\boldsymbol{\mathcal{L}}_\mathcal{H}\nu_{\Omega_h} - \nu_\Omega)\times(\boldsymbol{\mathcal{L}}_\mathcal{H}\boldsymbol{\mathcal{I}}_\mathcal{H}\varphi_\mathcal{H}) \mid \boldsymbol{\mathcal{L}}_\mathcal{V}\boldsymbol{\psi}_\mathcal{V}\big)_{L^2(\partial\Omega)^3} \\
&\quad + \big(\nu_\Omega\times((\boldsymbol{\mathcal{L}}_\mathcal{H}\boldsymbol{\mathcal{I}}_\mathcal{H} - \mathrm{Id})\varphi_\mathcal{H}) \mid \boldsymbol{\mathcal{L}}_\mathcal{V}\boldsymbol{\psi}_\mathcal{V}\big)_{L^2(\partial\Omega)^3}.
\end{aligned} \tag{5.69}$$

Here, ν_{Ω_h} and ν_Ω denote the outer unit normal vectors on Ω_h and Ω, respectively. Again, we consider these differences separately. For the first line, the Cauchy–Schwarz inequality together with the bound (2.3) for the cross product implies

$$\begin{aligned}&\int_{\partial\Omega} (\boldsymbol{\mathcal{L}}_\mathcal{H} \nu_{\Omega_h}) \times (\boldsymbol{\mathcal{L}}_\mathcal{H} \boldsymbol{\mathcal{I}}_\mathcal{H} \varphi_\mathcal{H})(\boldsymbol{\mathcal{L}}_\mathcal{V} \boldsymbol{\psi}_\mathcal{V})(|\det \mathrm{D}\,\mathcal{G}^{-1}| - 1)\,\mathrm{d}x \\ &\qquad \leq \|\boldsymbol{\mathcal{L}}_\mathcal{H} \nu_{\Omega_h}\|_{L^2(\partial\Omega)^3} \|\boldsymbol{\mathcal{L}}_\mathcal{H} \boldsymbol{\mathcal{I}}_\mathcal{H} \varphi_\mathcal{H}\|_{L^2(\partial\Omega)^3} \|\boldsymbol{\mathcal{L}}_\mathcal{V} \boldsymbol{\psi}_\mathcal{V}\|_{L^2(\partial\Omega)^3} \||\det \mathrm{D}\,\mathcal{G}^{-1}| - 1\|_{L^\infty(\partial\Omega)^3}.\end{aligned}$$

By construction the lift operators are also bounded operators on the boundary faces, i.e., we have $\boldsymbol{\mathcal{L}}_\mathcal{V}, \boldsymbol{\mathcal{L}}_\mathcal{H} \in \mathcal{L}(L^2(\partial\Omega_h)^3, L^2(\partial\Omega)^3)$, cf. [Elliott and Ranner, 2013, Prop. 4.13]. Thus, we get with

$$\|\boldsymbol{\mathcal{L}}_\mathcal{H} \nu_{\Omega_h}\|_{L^2(\partial\Omega)^3} \leq C \|\boldsymbol{\mathcal{L}}_\mathcal{H} \nu_{\Omega_h}\|_{L^\infty(\partial\Omega)^3} \leq C \|\nu_{\Omega_h}\|_{L^\infty(\partial\Omega_h)^3}$$

for a constant $C > 0$ depending on $\partial\Omega$ the estimate

$$\begin{aligned}&\int_{\partial\Omega} (\boldsymbol{\mathcal{L}}_\mathcal{H} \nu_{\Omega_h}) \times (\boldsymbol{\mathcal{L}}_\mathcal{H} \boldsymbol{\mathcal{I}}_\mathcal{H} \varphi_\mathcal{H})(\boldsymbol{\mathcal{L}}_\mathcal{V} \boldsymbol{\psi}_\mathcal{V})(|\det \mathrm{D}\,\mathcal{G}^{-1}| - 1)\,\mathrm{d}x \\ &\qquad \leq C \|\boldsymbol{\mathcal{L}}_\mathcal{H} \boldsymbol{\mathcal{I}}_\mathcal{H} \varphi_\mathcal{H}\|_{L^2(\partial\Omega)^3} \|\boldsymbol{\psi}_\mathcal{V}\|_{L^2(\partial\Omega)^3} \||\det \mathrm{D}\,\mathcal{G}^{-1}| - 1\|_{L^\infty(\partial\Omega)^3}.\end{aligned}$$

With the same arguments, we obtain for the second line of (5.69)

$$\begin{aligned}&\big((\boldsymbol{\mathcal{L}}_\mathcal{H} \nu_{\Omega_h} - \nu_\Omega) \times (\boldsymbol{\mathcal{L}}_\mathcal{H} \boldsymbol{\mathcal{I}}_\mathcal{H} \varphi_\mathcal{H}) \mid \boldsymbol{\mathcal{L}}_\mathcal{V} \boldsymbol{\psi}_\mathcal{V}\big)_{L^2(\partial\Omega)^3} \\ &\qquad \leq \|(\boldsymbol{\mathcal{L}}_\mathcal{H} \nu_{\Omega_h} - \nu_\Omega) \times (\boldsymbol{\mathcal{L}}_\mathcal{H} \boldsymbol{\mathcal{I}}_\mathcal{H} \varphi_\mathcal{H})\|_{L^2(\partial\Omega)^3} \|\boldsymbol{\mathcal{L}}_\mathcal{V} \boldsymbol{\psi}_\mathcal{V}\|_{L^2(\partial\Omega)^3} \\ &\qquad \leq \|\boldsymbol{\mathcal{L}}_\mathcal{H} \nu_{\Omega_h} - \nu_\Omega\|_{L^\infty(\partial\Omega)^3} \|\boldsymbol{\mathcal{I}}_\mathcal{H} \varphi_\mathcal{H}\|_{L^2(\partial\Omega)^3} \|\boldsymbol{\psi}_\mathcal{V}\|_{L^2(\partial\Omega)^3}.\end{aligned}$$

For the last line of (5.69), we obtain

$$\begin{aligned}&\big(\nu_\Omega \times ((\boldsymbol{\mathcal{L}}_\mathcal{H} \boldsymbol{\mathcal{I}}_\mathcal{H} - \mathrm{Id})\varphi_\mathcal{H}) \mid \boldsymbol{\mathcal{L}}_\mathcal{V} \boldsymbol{\psi}_\mathcal{V}\big)_{L^2(\partial\Omega)^3} \\ &\qquad \leq \|\nu_\Omega \times ((\boldsymbol{\mathcal{L}}_\mathcal{H} \boldsymbol{\mathcal{I}}_\mathcal{H} - \mathrm{Id})\varphi_\mathcal{H})\|_{L^2(\partial\Omega)^3} \|\boldsymbol{\mathcal{L}}_\mathcal{V} \boldsymbol{\psi}_\mathcal{V}\|_{L^2(\partial\Omega)^3} \\ &\qquad \leq \|\nu_\Omega\|_{L^\infty(\partial\Omega)^3} \|(\boldsymbol{\mathcal{L}}_\mathcal{H} \boldsymbol{\mathcal{I}}_\mathcal{H} - \mathrm{Id})\varphi_\mathcal{H}\|_{L^2(\partial\Omega)^3} \|\boldsymbol{\psi}_\mathcal{V}\|_{L^2(\partial\Omega_h)^3}.\end{aligned}$$

Using these bounds in (5.69) with $\|\boldsymbol{\mathcal{L}}_\mathcal{H} \nu_{\Omega_h}\|_{L^\infty(\partial\Omega)^3}, \|\nu_\Omega\|_{L^\infty(\partial\Omega)^3} \leq 1$ yields

$$\begin{aligned}\sum_{F_h \in \mathcal{F}_h^b} (\nu_{\Omega_h} \times \boldsymbol{\mathcal{I}}_\mathcal{H} \varphi_\mathcal{H} \mid \boldsymbol{\psi}_\mathcal{V})_{L^2(F_h)^3} &\leq \Big(\|(\boldsymbol{\mathcal{L}}_\mathcal{H} \boldsymbol{\mathcal{I}}_\mathcal{H} - \mathrm{Id})\varphi_\mathcal{H}\|_{L^2(\partial\Omega)^3} \\ &+ (\||\det \mathrm{D}\,\mathcal{G}^{-1}| - 1\|_{L^\infty(\partial\Omega)^3} + \|\boldsymbol{\mathcal{L}}_\mathcal{H} \nu_{\Omega_h} - \nu_\Omega\|_{L^\infty(\partial\Omega)^3}) \|\boldsymbol{\mathcal{L}}_\mathcal{H} \boldsymbol{\mathcal{I}}_\mathcal{H} \varphi_\mathcal{H}\|_{L^2(\partial\Omega)^3}\Big) \|\boldsymbol{\psi}_\mathcal{V}\|_{L^2(\partial\Omega_h)^3}.\end{aligned}$$

For the first term, the trace inequality (2.6) and the interpolation estimate (5.62) prove

$$\|(\boldsymbol{\mathcal{L}}_\mathcal{H} \boldsymbol{\mathcal{I}}_\mathcal{H} - \mathrm{Id})\varphi_\mathcal{H}\|_{L^2(\partial\Omega)^3} \leq C h^{r+\frac{1}{2}} \|\varphi_\mathcal{H}\|_{H^{r+1}(\Omega)^3}.$$

As for the volume terms, the same arguments as in [Hipp, 2017, Lem. 7.3] together with [Elliott and Ranner, 2013, Prop. 4.11] imply

$$\||\det \mathrm{D}\,\mathcal{G}^{-1}| - 1\|_{L^\infty(\partial\Omega)^3} \leq C h^{r+1}.$$

Furthermore, as $\partial\Omega$ is a C^{p+4}-boundary and due to the construction (5.50) of the approximation space $\boldsymbol{\mathcal{V}}_{\mathcal{H}}$ for the electric field with piecewise polynomials of degree $p+1$, we obtain from [Demlow, 2009, Prop. 2.3] the estimate

$$\|\boldsymbol{\mathcal{L}}_{\mathcal{H}}\nu_{\Omega_h} - \nu_\Omega\|_{L^\infty(\partial\Omega)^3} \leq Ch^{r+1}. \tag{5.70}$$

In addition, the construction of the lift $\boldsymbol{\mathcal{L}}$ and the interpolation operator $\boldsymbol{\mathcal{I}}$ as well as Sobolev's embedding (2.5) imply

$$\|\boldsymbol{\mathcal{L}}_{\mathcal{H}}\boldsymbol{\mathcal{I}}_{\mathcal{H}}\varphi_{\mathcal{H}}\|_{L^2(\partial\Omega)^3} \leq C\|\varphi_{\mathcal{H}}\|_{\mathcal{Y}_{\mathcal{H}}}.$$

Finally, since we obtain with the trace inequality (2.6) and the inverse estimate (4.1)

$$\|\boldsymbol{\psi}_{\mathcal{V}}\|_{L^2(\partial\Omega_h)^3} \leq Ch^{-\frac{1}{2}}\|\boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}},$$

collecting all results yields

$$\begin{aligned}(\nu_{\Omega_h} \times \boldsymbol{\mathcal{I}}_{\mathcal{H}}\varphi_{\mathcal{H}} \mid \boldsymbol{\psi}_{\mathcal{V}})_{L^2(F_h)^3} &\leq C\Big(h^{r+\frac{1}{2}}\|\varphi_{\mathcal{H}}\|_{H^{r+1}(\Omega)^3} + h^{r+1}\|\varphi_{\mathcal{H}}\|_{\mathcal{Y}_{\mathcal{H}}}\Big)h^{-\frac{1}{2}}\|\boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}} \\ &\leq Ch^r\|\varphi_{\mathcal{H}}\|_{H^{r+1}(\Omega)^3}\|\boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}}.\end{aligned} \tag{5.71}$$

Thus, using (5.64), (5.68), and (5.71) in (5.63) concludes the proof. □

Error estimate

In this section, we state wellposedness as well as an error estimate for the space discretization of the quasilinear Maxwell equations. It is based on the abstract result in Corollary 4.29 for local nonlinearities.

Theorem 5.13. *For $p \geq 3$, let $\Omega \subset \mathbb{R}^3$ be a bounded domain with C^{p+4}-boundary. Furthermore, let the assumptions of Theorem 3.5 be satisfied for $\boldsymbol{R} \in \big(0, (9\|\chi\|_{L^\infty(\Omega)})^{-\frac{1}{2}}\big)$ with $R < \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}\boldsymbol{R}$ and $m = p+2$. Then, there exists $h_0 > 0$ such that for all $h < h_0$, the solution $\boldsymbol{y} = (\boldsymbol{\mathfrak{H}}, \boldsymbol{\mathfrak{E}})$ of the discrete Maxwell equations with Kerr nonlinearity (5.57) satisfies*

$$\boldsymbol{\mathfrak{H}}, \boldsymbol{\mathfrak{E}} \in C^1(J_T, L^2(\Omega_h)^3) \cap C(J_T, B_{L^\infty(\Omega_h)^3}(\boldsymbol{R})).$$

Furthermore, for $t \in J_T$ we have

$$\|\mathcal{H}(t) - \boldsymbol{\mathcal{L}}_{\mathcal{V}}\boldsymbol{\mathfrak{H}}(t)\|_{L^2(\Omega)^3} + \|\mathcal{E}(t) - \boldsymbol{\mathcal{L}}_{\mathcal{H}}\boldsymbol{\mathfrak{E}}(t)\|_{L^2(\Omega)^3} \leq C_{\mathcal{H},\mathcal{E}}(1+t)\mathrm{e}^{Ct}h^p,$$

with constants $C_{\mathcal{H},\mathcal{E}}, C > 0$ independent of h, t, and T, but $C_{\mathcal{H},\mathcal{E}}$ depending on both fields $\mathcal{H}$ and $\mathcal{E}$ as well as the nonlinear susceptibility χ, including their derivatives.

Proof. We closely follow the arguments in the proof of Theorem 5.9, as the space discretization of the Maxwell equations is quite similar to the discretization of the multi-dimensional Westervelt equation.

Thus, we first determine $C_{\max}(h)$. For $\boldsymbol{C}_{\mathbf{A}}(h)$, we obtain from [Sturm, 2017, Thm. 3.14] the existence of a constant $C_{\mathcal{C}} > 0$ such that

$$\|\mathbf{A}_{\boldsymbol{\mathcal{V}}}\|_{\mathcal{L}(\boldsymbol{\mathcal{X}}_{\mathcal{H}},\boldsymbol{\mathcal{X}}_{\mathcal{V}})} \leq C_{\mathcal{C}}h^{-1}, \qquad \|\mathbf{A}_{\boldsymbol{\mathcal{H}}}\|_{\mathcal{L}(\boldsymbol{\mathcal{X}}_{\mathcal{V}},\boldsymbol{\mathcal{X}}_{\mathcal{H}})} \leq C_{\mathcal{C}}h^{-1}$$

holds. Hence, the definition (5.54) of $\mathbf{A}$ implies

$$\boldsymbol{C}_{\mathbf{A}}(h) = C_{\mathcal{C}} h^{-1}, \tag{5.72}$$

which finally together with Lemma 5.6 yields

$$C_{\max}(h) = C_{\text{inv},1} C_{\mathcal{C}} h^{-\frac{5}{2}}. \tag{5.73}$$

From Theorem 3.5, we obtain the existence of a solution $y = (\mathcal{H}, \mathcal{E})$ of the Maxwell equations with Kerr nonlinearity (3.23) such that

$$y \in C^1(J_T, H^{p+1}(\Omega)^6)$$

holds. Thus, we now check Assumption 4.28 for $\mathcal{Z} = \mathcal{Z}_{\mathcal{V}} \times \mathcal{Z}_{\mathcal{H}}$ given by (3.24).

(A_1) Due to $\boldsymbol{\mathcal{J}} = \boldsymbol{\mathcal{I}}$, (A_1) as a direct consequence of (5.62) since $p > \frac{5}{2}$.

(A_2) The definition (5.56) of the initial values together with $\boldsymbol{\mathcal{J}} = \boldsymbol{\mathcal{I}}$ directly yields $\boldsymbol{\mathcal{J}} y_0 = \boldsymbol{y}_0$. Thus, (A_2) is satisfied.

(A_3) Since we have $\boldsymbol{\mathcal{J}} = \boldsymbol{\mathcal{I}}$, (A_3) is trivially satisfied.

(A_5) As in [Hipp et al., 2019, Thm. 3.3], we obtain from (4.20), the Cauchy–Schwarz inequality, and (4.19) for $\xi \in \mathcal{Z}$ the estimate

$$\begin{aligned}
\|\mathcal{R}_{\mathrm{A}}\xi\|_{\boldsymbol{\mathcal{X}}} &= \sup_{\|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}}=1} ((\mathbf{A}\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{L}}_{\mathcal{X}}^{*}\mathrm{A})\xi \mid \boldsymbol{\xi})_{\boldsymbol{\mathcal{X}}} \\
&= \sup_{\|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}}=1} (\mathbf{A}\boldsymbol{\mathcal{I}}\xi \mid \boldsymbol{\xi})_{\boldsymbol{\mathcal{X}}} - (\mathrm{A}\xi \mid \boldsymbol{\mathcal{L}}\boldsymbol{\xi})_{\mathcal{X}} \\
&\leq \sup_{\|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}}=1} \Big((\mathbf{A}\boldsymbol{\mathcal{I}}\xi \mid \boldsymbol{\xi})_{\boldsymbol{\mathcal{X}}} - (\mathrm{A}\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}}\xi \mid \boldsymbol{\mathcal{L}}\boldsymbol{\xi})_{\mathcal{X}} \Big) + \sup_{\|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}}=1} (\mathrm{A}(\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}} - \mathrm{Id})\xi \mid \boldsymbol{\mathcal{L}}\boldsymbol{\xi})_{\mathcal{X}} \\
&\leq \sup_{\|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}}=1} \Big((\mathbf{A}\boldsymbol{\mathcal{I}}\xi \mid \boldsymbol{\xi})_{\boldsymbol{\mathcal{X}}} - (\mathrm{A}\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}}\xi \mid \boldsymbol{\mathcal{L}}\boldsymbol{\xi})_{\mathcal{X}} \Big) + \boldsymbol{C}_{\boldsymbol{\mathcal{L}}} \|\mathrm{A}(\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}} - \mathrm{Id})\xi\|_{\mathcal{X}}.
\end{aligned}$$

Thus, Lemma 5.12 and (5.67) together with Lemma 5.11 yield

$$\|\mathcal{R}_{\mathrm{A}}\xi\|_{\boldsymbol{\mathcal{X}}} \leq C h^p \|\xi\|_{H^{p+1}(\Omega)^3 \times H^{p+1}(\Omega)^3}. \tag{5.74}$$

Therefore, (A_5) is satisfied due to $p > \frac{5}{2}$.

(A_7) As Lemma 5.8 is also valid here, (A_7) follows from (5.43) since $p > \frac{5}{2}$.

Hence, Assumption 4.28 holds, so that Corollary 4.29 yields (5.44). Furthermore, we get from (4.61) due to $\boldsymbol{\mathcal{J}} = \boldsymbol{\mathcal{I}}$ and (5.56) with $y = (\mathcal{H}, \mathcal{E})$ and $\boldsymbol{y} = (\boldsymbol{\mathcal{H}}, \boldsymbol{\mathcal{E}})$ the estimate

$$\begin{aligned}
\|y(t) - \boldsymbol{\mathcal{L}}\boldsymbol{y}(t)\|_{\mathcal{X}} \leq C(1+t)\mathrm{e}^{Ct}\Big(&\sup_{[0,t]} \|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}})y\|_{\mathcal{X}} + \sup_{[0,t]} \|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}})\Lambda(y)\partial_t y\|_{\mathcal{X}} \\
&+ \sup_{[0,t]} \triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}\Lambda(y)\partial_t y) + \sup_{[0,t]} \|\mathcal{R}_{\mathrm{A}} y\|_{\boldsymbol{\mathcal{X}}} \Big).
\end{aligned}$$

As we further get from (5.43), (5.35), and (5.42) for $\xi \in \mathcal{Y}$ the bound

$$\begin{aligned}\triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}\xi) &\le Ch^p \|\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}}\xi\|_{\boldsymbol{\mathcal{X}}} \\ &\le Ch^p\|\xi\|_{\mathcal{Y}},\end{aligned} \tag{5.75}$$

we finally obtain with (5.74) and (5.62) the estimate

$$\begin{aligned}&\|\mathcal{H}(t) - \boldsymbol{\mathcal{L}}_{\mathcal{V}}\boldsymbol{\mathcal{H}}(t)\|_{L^2(\Omega)} + \|\mathcal{E}(t) - \boldsymbol{\mathcal{L}}_{\mathcal{V}}\boldsymbol{\mathcal{E}}(t)\|_{L^2(\Omega)} \\ &\quad\le C(1+t)\mathrm{e}^{Ct}h^p\Big(\sup_{[0,t]}\|\mathcal{H}\|_{H^{p+1}(\Omega)^3} + \sup_{[0,t]}\|\mathcal{E}\|_{H^{p+1}(\Omega)^3} + \sup_{[0,t]}\|\Lambda(y)\partial_t y\|_{H^p(\Omega)^3\times H^p(\Omega)^3}\Big).\end{aligned}$$

To conclude, since $H^p(\Omega)\times H^p(\Omega)$ is a Banach algebra, we obtain from the definition (3.22) of the nonlinearity due to the regularity of χ the estimate

$$\begin{aligned}&\sup_{[0,t]}\|\Lambda(y)\partial_t y\|_{H^p(\Omega)^3\times H^p(\Omega)^3} \\ &\quad\le C\Big(1+9\|\chi\|_{W^{p,\infty}(\Omega)}\sup_{[0,t]}\|\mathcal{E}\|^2_{H^p(\Omega)^3}\Big)\Big(\sup_{[0,t]}\|\partial_t\mathcal{H}\|_{H^p(\Omega)^3} + \sup_{[0,t]}\|\partial_t\mathcal{E}\|_{H^p(\Omega)^3}\Big),\end{aligned} \tag{5.76}$$

which yields the constant $C_{\mathcal{H},\mathcal{E}}$ and hence proves the result. □

Finally, we conclude this section with the following two remarks. First, we comment on the special construction of the approximation space $\boldsymbol{\mathcal{V}}_{\mathcal{H}}$ for the electric field.

Remark 5.14. *On the one hand, we emphasize that Theorem 5.13 is not valid if we define the approximation space for the electric field by $\boldsymbol{\mathcal{V}}_{\mathcal{H}} = \boldsymbol{\mathcal{V}}_{\mathcal{V}}$. In particular, in this case the approximation of the outer unit normal vector in (5.70) is only of order p, which drops the convergence to order $p-\frac{1}{2}$ in Lemma 5.12 and hence also in Theorem 5.13.*

On the other hand, increasing the polynomial degree for all elements $K_h \in \mathcal{T}_h$ seems to be an overly harsh measure, since [Demlow, 2009, Prop. 2.3] depends only on the approximation of the boundary. Instead, it would be sufficient to increase the polynomial degree only for elements near the boundary. However, as local increase of the polynomial degree conflicts with the construction of the interpolation operator $\boldsymbol{\mathcal{I}}_{\mathcal{V}} : C(\overline{\Omega}) \to C(\overline{\Omega_h})$, we do not investigate this further here.

In the following remark, we comment on the assumptions on the continuous solution.

Remark 5.15. *Theorem 5.13 is also valid for a bounded domain $\Omega \subset \mathbb{R}^3$ with C^{p+3}-boundary, as*

$$y \in C^1(J_T, H^p(\Omega)^6) \cap C(J_T, H^{p+1}(\Omega)^6)$$

is sufficient to prove the result.

However, since including this refined assumption in the abstract framework requires the introduction of an additional space $\widetilde{\mathcal{Z}}$ with $\mathcal{Z} \hookrightarrow \widetilde{\mathcal{Z}} \hookrightarrow \mathcal{Y}$ and

$$y \in C^1(J_T, \widetilde{\mathcal{Z}}) \cap C(J_T, \mathcal{Z}),$$

we refrain from this refinement in the abstract framework. Instead, Assumption 4.22 and Assumption 4.23 are based on $y \in C^1(J_T, \mathcal{Z})$, for the sake or presentation.

In addition, we emphasize that we do not require y to be globally regular as stated in (3.24). Instead, Theorem 5.13 is also valid for functions in $\mathcal{Z}$ being piecewise regular on every element of the exact triangulation $\mathcal{T}_h^{\mathrm{ex}}$.

CHAPTER 6

Time discretization of abstract problems

In the following we recall time-integration schemes for the full discretization of the abstract wave-type problem (3.3). This includes algebraically stable Runge–Kutta schemes and the leapfrog scheme. Furthermore, we briefly review recent results for algebraically stable Runge–Kutta schemes.

Throughout this chapter, let $\tau > 0$ be a constant time-step size and $T > 0$ be the final time. Moreover, let $N \in \mathbb{N}$ with $N\tau \leq T$ be the number of time steps. Finally, we employ for $n = \{0, \dots, N\}$ the notation $t_n = n\tau$.

6.1 Algebraically stable, coercive Runge–Kutta schemes

In this section we recall algebraically stable, coercive Runge–Kutta schemes with a special focus on the implicit midpoint rule. Additionally, we briefly review the main results from [Hochbruck and Pažur, 2017], [Hochbruck et al., 2018], and [Kovács and Lubich, 2018].

We consider abstract ordinary differential equations of the form

$$\begin{cases} y'(t) = f\big(t, y(t)\big), & t \in J_T, \\ y(0) = y_0, \end{cases} \tag{6.1}$$

where $J_T = [0, T]$ denotes the time interval, $y_0 \in U$ is the initial value and $f : J_T \times U \to \mathbb{R}$, for an open, simply-connected set $U \subset \mathbb{R}^d$ with $d \in \mathbb{N}$.

For $s \in \mathbb{N}$ and $i, j = 1, \dots, s$, let $b_i, a_{ij} \in \mathbb{R}$ and $c_i = \sum_{j=1}^{s} a_{ij}$. Then, the s-stage Runge–Kutta scheme corresponding to these coefficients is given by

$$\begin{aligned} Y'_{ni} &= f(t_n + c_i\tau, Y_{ni}), && i = 1, \dots, s, \\ Y_{ni} &= y_n + \tau \sum_{j=1}^{s} a_{ij} Y'_{nj}, && i = 1, \dots, s, \\ y_{n+1} &= y_n + \tau \sum_{i=1}^{s} b_i Y'_{ni}, \end{aligned}$$

for $n = 0, \dots, N-1$. Note that the iterates y_n, Y_{ni}, and Y'_{ni} approximate the point evaluations $y(t_n)$, $y(t_n + c_i\tau)$, and $y'(t_n + c_i\tau)$ of the exact solution, respectively.

An important tool for the classification of Runge–Kutta schemes are the different concepts of stability. Hence, in the following we recall the concepts of A-stability, B-stability, and algebraic stability.

We first consider for $\mu \in \mathbb{C}$ the test problem

$$\begin{cases} y'(t) = \mu y, \quad t \in J_T, \\ y(0) = y_0. \end{cases} \tag{6.2}$$

Note that the solution $y(t) = y_0 \mathrm{e}^{\mu t}$ of (6.2) is bounded for all $\mu \in \mathbb{C}$ with $\operatorname{Re}\mu \leq 0$. This motivates the following definition of A-stability.

Definition 6.1. *A Runge–Kutta scheme with constant time-step size $\tau > 0$ is called A-stable if the numerical approximations $\{y_n\}_{n\in\mathbb{N}}$ obtained by the application of the scheme to the initial value problem (6.2) are bounded for all $\operatorname{Re}\mu \leq 0$.*

Next, we consider (6.1) with $f : J_T \times U \to \mathbb{R}$ satisfying the additional condition

$$\operatorname{Re}\left(f(t,y) - f(t,z) \mid y - z\right)_* \leq 0, \qquad t \in J_T,\ y, z \in U, \tag{6.3}$$

where $(\cdot \mid \cdot)_*$ is an inner product on $\mathbb{R}^d$ with corresponding norm $\|\cdot\|_*$. For initial values $y_0, z_0 \in U$ with corresponding solutions $y, z : J_T \to U$, we then have

$$\|y(t) - z(t)\|_* \leq \|y_0 - z_0\|_*, \qquad t \in J_T.$$

This motivates the following definition of B-stability.

Definition 6.2. *A Runge–Kutta scheme is called B-stable if the contractivity condition (6.3) implies for all time-step sizes $\tau > 0$ the relation*

$$\|y_1 - z_1\|_* \leq \|y_0 - z_0\|_*.$$

Here, $y_1, z_1 \in U$ are the approximations obtained after one step of the scheme applied to (6.1) with initial values $y_0, z_0 \in U$, respectively.

Since the contractivity condition (6.3) is satisfied for all test problems (6.2) with $\operatorname{Re}\mu \leq 0$, A-stability is implied by B-stability.

Definition 6.3. *An s-stage Runge–Kutta scheme is called algebraically stable if the coefficients satisfy the following conditions.*

(i) For all $i = 1, \dots, s$, we have $b_i \geq 0$.

(ii) The matrix $(m_{ij})_{i,j=1,\dots,s}$ given by $m_{ij} = b_i a_{ij} + b_j a_{ji} - b_i b_j$ is positive semidefinite.

Algebraic stability implies B-stability, but has the advantage of being easy to check. Note that Gauß and Radau collocation methods are algebraically stable. Moreover, these methods satisfy the following coercivity condition.

Definition 6.4. *An s-stage Runge–Kutta scheme with Runge–Kutta matrix $\mathcal{O} = (a_{ij})_{i,j=1,\dots,s}$ is called coercive if there exist $\alpha > 0$ and a diagonal, positive definite matrix $\mathcal{D} \in \mathbb{R}^{s\times s}$ such that*

$$\xi^T \mathcal{D}\mathcal{O}^{-1}\xi \geq \alpha \xi^T \mathcal{D}\xi, \qquad \xi \in \mathbb{R}^s.$$

Based on these concepts, we now review some recent results on the time discretization of the quasilinear evolution equation (3.3).

In [Hochbruck and Pažur, 2017] and [Hochbruck et al., 2018], the authors investigate the time discretization of (3.3) with right-hand side

$$\mathrm{F}(\cdot,\xi) = Q(\xi)\xi, \qquad \xi \in B_{\mathcal{Y}}(R),$$

for a given nonlinearity Q. In particular, the quasilinear Maxwell equations (3.16) on the full space $\mathbb{R}^3$ as well as both the quasilinear Maxwell equations (3.16) and the quasilinear wave equation (3.10) on bounded domains $\Omega \subset \mathbb{R}^3$, subject to Dirichlet boundary conditions, are considered as specific examples.

Based on the linear operators

$$\widetilde{\mathcal{A}}(\xi) \coloneqq \Lambda(\xi)^{-1}(\mathrm{A} + Q(\xi)), \qquad \xi \in B_{\mathcal{Y}}(R),$$

the semi-implicit Euler method

$$y_{n+1} = y_n + \tau\widetilde{\mathcal{A}}(y_n)y_{n+1}, \qquad n = 0,\dots,N-1,$$

as well as the implicit Euler method

$$y_{n+1} = y_n + \tau\widetilde{\mathcal{A}}(y_{n+1})y_{n+1}, \qquad n = 0,\dots,N-1,$$

are analyzed in [Hochbruck and Pažur, 2017]. More precisely, based on the wellposedness result presented in [Müller, 2014, Thm. 3.41] for the continuous problem, the authors prove wellposedness and error estimates in the $\mathcal{X}$- and $\mathcal{Y}$-norm. In [Hochbruck et al., 2018], the wellposedness analysis and the error analysis are extended to the general class of algebraically stable, coercive Runge–Kutta schemes.

In [Kovács and Lubich, 2018], quasilinear evolution equations of the form

$$\partial_t y(t) = \widehat{\mathcal{A}}(y(t))y(t) + \mathcal{F}(y(t)), \qquad t \in J_T,$$

are considered with a linear differential operator $\widehat{\mathcal{A}}(\xi)$, $\xi \in B_{\mathcal{Y}}(R)$, and a regular function $\mathcal{F}$. As shown in [Kato, 1975], this framework includes various problems posed on the full space, e.g., the quasilinear wave equation and the quasilinear Maxwell equations.

In particular, two variants of the implicit midpoint rule

$$y_{n+1} = y_n + \tau\widehat{\mathcal{A}}(\underline{y}_{n+1/2})y_{n+1/2} + \tau\mathcal{F}(y_{n+1/2}), \qquad n = 0,\dots,N-1, \tag{6.4}$$

with

$$y_{n+1/2} = \frac{y_{n+1} + y_n}{2}, \qquad n = 0,\dots,N-1, \tag{6.5}$$

are considered in [Kovács and Lubich, 2018, Sec. 3]. On the one hand, we obtain from (6.4) for $\underline{y}_{n+1/2} = y_{n+1/2}$ the fully implicit midpoint rule. On the other hand, for

$$\underline{y}_{n+1/2} = \frac{3y_n - y_{n-1}}{2}, \qquad n = 1, \dots, N-1, \tag{6.6}$$

and $\underline{y}_{1/2} = y_0$, (6.4) corresponds to the linearly implicit midpoint rule. For both variants, the authors prove wellposedness and an error estimate in $\mathcal{Y}$. Moreover, these results are extended to algebraically stable, coercive Runge–Kutta schemes, including error estimates in $\mathcal{Y}$ with both the stage order plus 1 and the classical order for sufficiently regular solutions.

To conclude this section, we discuss the extendibility of the techniques used to prove these results for the time discretization of quasilinear wave-type problems with respect to the full discretization.

Remark 6.5. *An essential requirement for the proof of the previously discussed results for the time discretization of quasilinear wave-type problems is the existence of a continuous isomorphism $S : \mathcal{Y} \to \mathcal{X}$ such that*

$$S\mathcal{A}(\xi)S^{-1} = \mathcal{A}(\xi) + \widetilde{\mathcal{B}}(\xi), \qquad \xi \in B_{\mathcal{Y}}(R), \tag{6.7}$$

holds, with $\widetilde{\mathcal{B}}(\xi) \in \mathcal{L}(\mathcal{X})$ being uniformly bounded. This condition arises because the analysis essentially employs the semigroup theory presented in Section 4.2.2. In particular, (6.7) corresponds to assumption (H_2) in Theorem 4.14.

For (3.3) considered on bounded spatial domains, (6.7) is not only restrictive due to the associated regularity assumptions on the boundary, but for specific boundary conditions also with respect to the additional assumptions on the nonlinearities, cf. [Müller, 2014, Sec. 4.1], where quasilinear Maxwell equations with perfectly conducting boundary conditions are investigated.

As discussed in Remark 4.17, up to our knowledge it is not clear whether such isomorphisms exist in the spatially discrete setting. Thus, we analyze the full discretization of quasilinear wave-type equations based on the same techniques as used in the analysis of the space discretization in Chapter 4, i.e., we use the inverse estimates (4.1) instead of (6.7).

Next, we focus on the leapfrog scheme.

6.2 Leapfrog scheme

In this section we consider the one-step formulation of the leapfrog scheme, which is an explicit time-integration scheme for partitioned systems.

For $d \in \mathbb{N}$, let $f, g : \mathbb{R}^d \to \mathbb{R}^d$ be sufficiently smooth functions and $u_0, v_0 \in \mathbb{R}^d$. In the following, we consider the system

$$\begin{aligned} \partial_t u &= f(v), \quad u(0) = u_0, \\ \partial_t v &= g(u), \quad v(0) = v_0. \end{aligned} \tag{6.8}$$

The leapfrog scheme applied to (6.8) is given by

$$\widehat{u}_{n+1/2} = u_n + \tfrac{\tau}{2} f(v_n), \tag{6.9a}$$

$$v_{n+1} = v_n + \tau g(\widehat{u}_{n+1/2}), \tag{6.9b}$$

$$u_{n+1} = \widehat{u}_{n+1/2} + \tfrac{\tau}{2} f(v_{n+1}), \tag{6.9c}$$

for $n = 0, \dots, N-1$. Note that this is an explicit scheme, which is known to be of classical order 2. Furthermore, except for the first time step it requires only a single evaluation of the right-hand sides f and g, as the evaluation of f in (6.9a) is available from the previous time step.

To apply the leapfrog scheme, we assume additionally to Assumption 3.1 that the problem (3.3) is a coupled system of partial differential equations. More precisely, we assume for spaces $\mathcal{X} = \mathcal{X}_{\mathcal{V}} \times \mathcal{X}_{\mathcal{H}}$ and $\mathcal{Y} = \mathcal{Y}_{\mathcal{V}} \times \mathcal{Y}_{\mathcal{H}}$, that the quasilinear evolution equation (3.1) is of the form

$$\begin{cases} \Lambda(y(t))\partial_t y(t) = \mathrm{A}y(t) + \mathrm{F}(t, y(t)), \quad y = \begin{pmatrix} u \\ v \end{pmatrix}, t \in J_T, \\ \qquad\qquad y(0) = y_0, \end{cases} \tag{6.10}$$

with

$$\Lambda(y) = \begin{pmatrix} \Lambda_{\mathcal{V}}(u,v) & 0 \\ 0 & \Lambda_{\mathcal{H}}(u,v) \end{pmatrix}, \qquad \mathrm{A} = \begin{pmatrix} 0 & \mathrm{A}_{\mathcal{V}} \\ \mathrm{A}_{\mathcal{H}} & 0 \end{pmatrix}, \qquad \mathrm{F}(t,y) = \begin{pmatrix} \mathrm{F}_{\mathcal{V}}(t,u,v) \\ \mathrm{F}_{\mathcal{H}}(t,u,v) \end{pmatrix}.$$

As in (3.4), we define

$$\mathcal{A}(y) = \begin{pmatrix} 0 & \mathcal{A}_{\mathcal{V}}(u,v) \\ \mathcal{A}_{\mathcal{H}}(u,v) & 0 \end{pmatrix} := \begin{pmatrix} 0 & \Lambda_{\mathcal{V}}(u,v)^{-1}\mathrm{A}_{\mathcal{V}} \\ \Lambda_{\mathcal{H}}(u,v)^{-1}\mathrm{A}_{\mathcal{H}} & 0 \end{pmatrix}, \quad y = \begin{pmatrix} u \\ v \end{pmatrix} \in B_{\mathcal{Y}}(R),$$

and

$$\mathcal{F}(t,y) = \begin{pmatrix} \mathcal{F}_{\mathcal{V}}(t,u,v) \\ \mathcal{F}_{\mathcal{H}}(t,u,v) \end{pmatrix} := \begin{pmatrix} \Lambda_{\mathcal{V}}(u,v)^{-1}\mathrm{F}_{\mathcal{V}}(t,u,v) \\ \Lambda_{\mathcal{H}}(u,v)^{-1}\mathrm{F}_{\mathcal{H}}(t,u,v) \end{pmatrix}, \qquad t \in J_T,\ y = \begin{pmatrix} u \\ v \end{pmatrix} \in B_{\mathcal{Y}}(R).$$

We emphasize that these assumptions are satisfied for the specific examples, cf. Section 3.2 and Section 3.3. In particular, for the wave equation, we seek the position u and the velocity $v = \partial_t u$, whereas for the Maxwell equations, the quantities of interest are the magnetic field $u = \mathcal{H}$ and the electric field $v = \mathcal{E}$.

Note that the abstract quasilinear system (6.10) does not directly fit in the framework (6.8), as the right-hand sides depend on both unknowns. Nevertheless, we present a modified version of the leapfrog scheme for the full discretization of (6.10) in Section 7.2.

CHAPTER 7

Full discretization of abstract problems

We now consider the full discretization of the quasilinear evolution equation (3.3). Following the method-of-lines approach, we combine the space discretization from Chapter 4 with the time-integration schemes presented in Chapter 6. In particular, we prove for the implicit midpoint rules and the leapfrog scheme wellposedness as well as a rigorous error estimate.

Throughout this chapter, we employ the following short notation for the sake of presentation. For the numerical solution $\widetilde{y}_n = y(t_n)$ of (3.3) at time t_n, we write

$$\widetilde{\mathcal{A}}_n := \mathcal{A}(\widetilde{y}_n), \qquad \widetilde{\mathcal{F}}_n := \mathcal{F}(t_n, \widetilde{y}_n), \qquad \widetilde{\boldsymbol{\mathcal{A}}}_n := \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}\widetilde{y}_n), \qquad \widetilde{\boldsymbol{\mathcal{F}}}_n := \boldsymbol{\mathcal{F}}(t_n, \boldsymbol{\mathcal{I}}\widetilde{y}_n). \tag{7.1}$$

Correspondingly, for the numerical approximation $\boldsymbol{y}_n$ obtained by either of the time-integration schemes applied to (4.12), we set

$$\boldsymbol{\mathcal{A}}_n := \boldsymbol{\mathcal{A}}(\boldsymbol{y}_n), \qquad\qquad \boldsymbol{\mathcal{F}}_n := \boldsymbol{\mathcal{F}}(t_n, \boldsymbol{y}_n). \tag{7.2}$$

7.1 Linearly and fully implicit midpoint rule

In this section we consider the fully implicit midpoint rule applied to (3.3)

$$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + \tau \boldsymbol{\mathcal{A}}_{n+1/2} \boldsymbol{y}_{n+1/2} + \tau \boldsymbol{\mathcal{F}}_{n+1/2}, \tag{7.3}$$

with

$$\boldsymbol{y}_{n+1/2} = \frac{\boldsymbol{y}_{n+1} + \boldsymbol{y}_n}{2}, \tag{7.4}$$

cf. (6.4) and (6.5). As proposed in [Kovács and Lubich, 2018] for quasilinear wave-type problems on unbounded domains, we also consider the linearly implicit midpoint rule

$$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + \tau \underline{\boldsymbol{\mathcal{A}}}_{n+1/2} \boldsymbol{y}_{n+1/2} + \tau \underline{\boldsymbol{\mathcal{F}}}_{n+1/2}, \tag{7.5}$$

where the approximations at the midpoints appearing in the nonlinearities are replaced by

$$\underline{\boldsymbol{y}}_{n+1/2} = \frac{3\boldsymbol{y}_n - \boldsymbol{y}_{n-1}}{2}, \qquad n = 1, \ldots, N-1, \tag{7.6}$$

and $\underline{\boldsymbol{y}}_{1/2} = \boldsymbol{y}_0$, cf. (6.4) and (6.6). This can be seen as an extrapolation based on the approximations computed at two previous time steps. The scheme is computationally efficient, as the implementation only requires the solution of one linear system in every step, whereas we have to solve a nonlinear system of equations in every step of the fully implicit scheme (7.3). We will also use it for the analysis of the fully implicit scheme, by approximating the nonlinear scheme by a sequence of linear ones.

Remark 7.1. *All results can also be generalized to variable step sizes $\tau_i \in [\tau_{\min}, \tau_{\max}]$, $i = 1, \ldots, N$, for $0 < \tau_{\min} < \tau_{\max} < \infty$ fixed. For the linearly implicit midpoint rule, we then use the extrapolations*

$$\underline{\boldsymbol{y}}_{n+1/2} = \boldsymbol{y}_n + \frac{\tau_{n+1}}{2\tau_n}(\boldsymbol{y}_n - \boldsymbol{y}_{n-1}), \qquad n = 1, \ldots, N-1.$$

Proving the existence of a unique solution of (7.3) in the required spaces is significantly more involved as its counterpart for the linearized scheme (7.5). Hence, we focus in the first part on the linearly implicit midpoint rule and extend these results in the second part to the fully implicit midpoint rule.

We first derive error recursions for both schemes. Hence, let $n = 0, \ldots, N-1$. We first state that the exact solution y of (3.3) satisfies a perturbed version of (7.3) and (7.5). To do so, we introduce for $\vartheta \in [0, 1]$ the notation

$$\widetilde{y}_{n+\vartheta} = y\big((1-\vartheta)t_n + \vartheta t_{n+1}\big).$$

Hence, we get

$$\widetilde{y}_{n+1} = \widetilde{y}_n + \tau \widetilde{\mathcal{A}}_{n+1/2}\widetilde{y}_{n+1/2} + \tau \widetilde{\mathcal{F}}_{n+1/2} + \delta_{n+1}, \tag{7.7}$$

for some defects δ_{n+1}. We further employ the defects

$$\widehat{\delta}_{n+1/2} = \widetilde{y}_{n+1/2} - \frac{\widetilde{y}_{n+1} + \widetilde{y}_n}{2} \tag{7.8}$$

and the discrete errors

$$\boldsymbol{e}_n := \boldsymbol{\mathcal{J}}\widetilde{y}_n - \boldsymbol{y}_n, \qquad \boldsymbol{e}_{n+1/2} := \frac{\boldsymbol{e}_{n+1} + \boldsymbol{e}_n}{2}. \tag{7.9}$$

These errors satisfy

$$\boldsymbol{e}_{n+1} = \boldsymbol{e}_n + \tau \widetilde{\boldsymbol{\mathcal{A}}}_{n+1/2}\boldsymbol{e}_{n+1/2} + \tau \boldsymbol{g}_{n+1}, \tag{7.10}$$

where the right-hand side is for the fully implicit midpoint rule given by $\boldsymbol{g}_{n+1} = \boldsymbol{g}_{n+1}^{\mathrm{FI}}$, with

$$\boldsymbol{g}_{n+1}^{\mathrm{FI}} = (\widetilde{\boldsymbol{\mathcal{A}}}_{n+1/2} - \boldsymbol{\mathcal{A}}_{n+1/2})\boldsymbol{y}_{n+1/2} - \boldsymbol{\mathcal{F}}_{n+1/2} + \tfrac{1}{\tau}\boldsymbol{\mathcal{J}}(\widetilde{y}_{n+1} - \widetilde{y}_n) + \widetilde{\boldsymbol{\mathcal{A}}}_{n+1/2}\boldsymbol{\mathcal{J}}(\widehat{\delta}_{n+1/2} - \widetilde{y}_{n+1/2}).$$

Similarly to the spatially discrete case, we apply the adjoint of the lift operator $\boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{n+1/2}]$ with respect to the weighted inner products to (7.7) and add the result to the right-hand side, cf. (4.43). This yields

$$\begin{aligned}\boldsymbol{g}_{n+1}^{\mathrm{FI}} = (\widetilde{\boldsymbol{\mathcal{A}}}_{n+1/2} - \boldsymbol{\mathcal{A}}_{n+1/2})\boldsymbol{y}_{n+1/2} + \widetilde{\boldsymbol{\mathcal{F}}}_{n+1/2} - \boldsymbol{\mathcal{F}}_{n+1/2} + \tfrac{1}{\tau}(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{n+1/2}])(\widetilde{y}_{n+1} - \widetilde{y}_n)\\ + \boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{n+1/2}]\widetilde{\mathcal{F}}_{n+1/2} - \widetilde{\boldsymbol{\mathcal{F}}}_{n+1/2} + (\widetilde{\boldsymbol{\mathcal{A}}}_{n+1/2}\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{n+1/2}]\widetilde{\mathcal{A}}_{n+1/2})(\widehat{\delta}_{n+1/2} - \widetilde{y}_{n+1/2})\\ + \boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{n+1/2}]\widetilde{\mathcal{A}}_{n+1/2}\widehat{\delta}_{n+1/2} + \tfrac{1}{\tau}\boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{n+1/2}]\delta_{n+1}.\end{aligned} \tag{7.11}$$

For the linearly implicit midpoint rule, the same approach yields (7.10), with $\boldsymbol{g}_{n+1} = \boldsymbol{g}_{n+1}^{\mathrm{LI}}$ and

$$\begin{aligned}\boldsymbol{g}_{n+1}^{\mathrm{LI}} = (\widetilde{\boldsymbol{\mathcal{A}}}_{n+1/2} - \underline{\boldsymbol{\mathcal{A}}}_{n+1/2})\boldsymbol{y}_{n+1/2} + \widetilde{\boldsymbol{\mathcal{F}}}_{n+1/2} - \underline{\boldsymbol{\mathcal{F}}}_{n+1/2} + \tfrac{1}{\tau}(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{n+1/2}])(\widetilde{y}_{n+1} - \widetilde{y}_n)\\ + \boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{n+1/2}]\widetilde{\mathcal{F}}_{n+1/2} - \widetilde{\boldsymbol{\mathcal{F}}}_{n+1/2} + (\widetilde{\boldsymbol{\mathcal{A}}}_{n+1/2}\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{n+1/2}]\widetilde{\mathcal{A}}_{n+1/2})(\widehat{\delta}_{n+1/2} - \widetilde{y}_{n+1/2})\\ + \boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{n+1/2}]\widetilde{\mathcal{A}}_{n+1/2}\widehat{\delta}_{n+1/2} + \tfrac{1}{\tau}\boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{n+1/2}]\delta_{n+1},\end{aligned} \tag{7.12}$$

which only linearly depends on $\boldsymbol{y}_{n+1/2}$ and hence also on the unknown $\boldsymbol{y}_{n+1}$. Here, we used the short notation

$$\underline{\boldsymbol{\mathcal{A}}}_{n+1/2} \coloneqq \boldsymbol{\mathcal{A}}(\underline{\boldsymbol{y}}_{n+1/2}), \qquad \underline{\boldsymbol{\mathcal{F}}}_{n+1/2} \coloneqq \boldsymbol{\mathcal{F}}(t_{n+1/2}, \underline{\boldsymbol{y}}_{n+1/2}). \tag{7.13}$$

We introduce the following sharper version of Assumption 4.23, to gather all assumptions on the solution of (3.3) in one statement.

Assumption 7.2. *Let Assumption 3.1 be satisfied. The quasilinear Cauchy problem (3.3) has a unique solution with maximal time of existence $t^*(y_0) > 0$, i.e., for every $T < t^*(y_0)$ there is a unique solution y of (3.3) satisfying*

$$y \in C^3(J_T, \mathcal{X}) \cap C^2(J_T, \mathcal{Y}) \cap C^1(J_T, \mathcal{Z}) \cap C(J_T, B_{\mathcal{Y}}(R)).$$

Additionally, there are $R^{\partial_t}, R^{\mathcal{A}} > 0$ such that the solution satisfies

$$\|\partial_t y(t)\|_{\mathcal{Y}} < R^{\partial_t}, \qquad \|\mathcal{A}(y(t))y(t)\|_{\mathcal{Y}} < R^{\mathcal{A}}$$

uniformly for $t \in J_T$.

We now state the bound for the discrete error of the full discretization with either of the implicit midpoint rules. To do so, we require that the discretization parameters $\tau, h > 0$ are chosen such that there exist $\varepsilon_0, C_0 > 0$ with

$$\tau C_{\max}(h)^{\frac{1}{2}} \le C_0 h^{\varepsilon_0}, \tag{7.14}$$

where $C_{\max}(h)$ is the constant defined in (4.45). Despite the fact that the implicit midpoint rule is in general unconditionally stable when applied to linear problems, we can not avoid this step size restriction here, as it is necessary to bound the iterates in the stronger space $\boldsymbol{\mathcal{Y}}$ for the scheme to be wellposed. Hence, as stated in [Makridakis, 1993], this restriction is not induced by the techniques used for the analysis, but inherent in the problem itself.

Theorem 7.3. *Let Assumption 4.22 as well as Assumption 7.2 be satisfied and $T < t^*(y_0)$. Then, there exist $h_0, \tau_0 > 0$ such that for all $h < h_0$ and $\tau < \tau_0$ satisfying the step size restriction (7.14), both the fully and the linearly implicit midpoint rule (7.3) and (7.5), respectively, are wellposed and satisfy for $n = 0, \dots, N$ the error estimate*

$$\begin{aligned} \|\boldsymbol{e}_n\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y_n})} \leq C(1+t_n)\mathrm{e}^{Ct_n}\Big(&\|\boldsymbol{e}_0\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y_0})} + \tau^2\big(\sup_{[0,t_n]}\|\partial_t y\|_{\mathcal{Y}} + \sup_{[0,t_n]}\|\partial_t^2 y\|_{\mathcal{Y}} + \sup_{[0,t_n]}\|\partial_t^3 y\|_{\mathcal{X}}\big) \\ &+ \sup_{[0,t_n]}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})y\|_{\boldsymbol{\mathcal{X}}} + \sup_{s_1,s_2\in[0,t_n]}\|\mathcal{R}_\Lambda(y(s_1))\partial_t y(s_2)\|_{\boldsymbol{\mathcal{X}}} \\ &+ \sup_{[0,t_n]}\|\mathcal{R}_\mathrm{A} y\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_n]}\|\mathcal{R}_\mathrm{F}(\cdot, y)\|_{\boldsymbol{\mathcal{X}}}\Big), \end{aligned} \tag{7.15}$$

with a constant $C > 0$ independent of τ, h, n and T. Moreover, τ_0 depends on the constants ε_0, C_0 from the step size restriction (7.14).

The proof of this theorem is postponed to the respective subsection.

The following corollary, which is a direct consequence of the theorem, then yields convergence of the full discretization with either of the implicit midpoint rules.

Corollary 7.4. *Under the assumptions of Theorem 7.3, the error of the linearly or fully implicit midpoint rule (7.3) or (7.5), respectively, for $n = 0, \dots, N$ is bounded by*

$$\begin{aligned} \|y(t_n) - \boldsymbol{\mathcal{L}}\boldsymbol{y}_n\|_{\mathcal{X}} \leq \|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{J}})y(t_n)\|_{\mathcal{X}} + C(1+t_n)\mathrm{e}^{Ct_n}\Big(&\|\boldsymbol{\mathcal{J}}y_0 - \boldsymbol{y}_0\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_n]}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})y\|_{\boldsymbol{\mathcal{X}}} \\ &+ \tau^2\big(\sup_{[0,t_n]}\|\partial_t y\|_{\mathcal{Y}} + \sup_{[0,t_n]}\|\partial_t^2 y\|_{\mathcal{Y}} + \sup_{[0,t_n]}\|\partial_t^3 y\|_{\mathcal{X}}\big) \\ &+ \sup_{s_1,s_2\in[0,t_n]}\|\mathcal{R}_\Lambda(y(s_1))\partial_t y(s_2)\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_n]}\|\mathcal{R}_\mathrm{A} y\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_n]}\|\mathcal{R}_\mathrm{F}(\cdot, y)\|_{\boldsymbol{\mathcal{X}}}\Big), \end{aligned} \tag{7.16}$$

with a constant $C > 0$ independent of τ, h, n and T. Furthermore, we have

$$\|y(t_n) - \boldsymbol{\mathcal{L}}\boldsymbol{y}_n\|_{\mathcal{X}} \to 0, \qquad n = 0, \dots, N, \tag{7.17}$$

for $\tau, h \to 0$ satisfying the step size restriction (7.14).

Proof. As in Theorem 4.20, we first split the error into

$$\|y(t_n) - \boldsymbol{\mathcal{L}}\boldsymbol{y}_n\|_{\mathcal{X}} \leq \|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{J}})y(t_n)\|_{\mathcal{X}} + \boldsymbol{C}_{\boldsymbol{\mathcal{L}}}\|\boldsymbol{e}_n\|_{\boldsymbol{\mathcal{X}}},$$

where the first term already appears in the right-hand side of (7.16). The bound for the second term follows directly from the norm equivalence (4.10) and the bound for the discrete error (7.15). Finally, Assumption 4.22 yields (7.17). □

Furthermore, we combine the results from Section 4.4 for the discretization of local nonlinearities with the full discretization, which is in particular useful for the later examples.

Corollary 7.5. *Let Assumption 4.26 be satisfied. Then, the statements of Theorem 7.3 and Corollary 7.4 are also valid if we replace Assumption 4.22 by Assumption 4.28. In particular, the error of the linearly or fully implicit midpoint rule (7.3) or (7.5), respectively, for $n = 0, \dots, N$ is bounded by*

$$\begin{aligned}\|y(t_n) - \boldsymbol{\mathcal{L}} \boldsymbol{y}_n\|_{\mathcal{X}} \leq \|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{J}})y(t_n)\|_{\mathcal{X}} + C(1+t_n)\mathrm{e}^{Ct_n}\Big(&\|\boldsymbol{\mathcal{J}}y_0 - \boldsymbol{y}_0\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_n]}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})y\|_{\boldsymbol{\mathcal{X}}} \\ &+ \tau^2\big(\sup_{[0,t_n]}\|\partial_t y\|_{\mathcal{Y}} + \sup_{[0,t_n]}\|\partial_t^2 y\|_{\mathcal{Y}} + \sup_{[0,t_n]}\|\partial_t^3 y\|_{\mathcal{X}}\big) + \sup_{[0,t_n]}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\partial_t y\|_{\boldsymbol{\mathcal{X}}} \\ &+ \sup_{s_1,s_2\in[0,t_n]}\|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}})\Lambda(y(s_1))\partial_t y(s_2)\|_{\mathcal{X}} + \sup_{s_1,s_2\in[0,t_n]}\triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}\Lambda(y(s_1))\partial_t y(s_2)) \\ &+ \sup_{[0,t_n]}\|\mathcal{R}_{\mathrm{A}} y\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_n]}\|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}})\mathrm{F}(\cdot,y)\|_{\mathcal{X}} + \sup_{[0,t_n]}\triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}\mathrm{F}(\cdot,y))\Big),\end{aligned}$$

with a constant $C > 0$ independent of τ, h, n and T.

Proof. The result is a direct consequence of Lemma 4.27 and Theorem 7.3. □

We conclude this section with a bound for the defects δ_{n+1} and $\widehat{\delta}_{n+1/2}$, as these defects arise in the analysis of both schemes.

Lemma 7.6. *Let Assumption 7.2 be true. Then, the defects satisfy*

$$\|\tfrac{1}{\tau}\delta_{n+1}\|_{\mathcal{X}} \leq C\tau^2 \sup_{[t_n,t_{n+1}]}\|\partial_t^3 y\|_{\mathcal{X}}, \qquad \|\widehat{\delta}_{n+1/2}\|_{\mathcal{Y}} \leq C\tau^2 \sup_{[t_n,t_{n+1}]}\|\partial_t^2 y\|_{\mathcal{Y}}$$

for $n = 0, \dots, N-1$.

Proof. Using (3.3) in (7.7) implies

$$\delta_{n+1} = \widetilde{y}_{n+1} - \widetilde{y}_n - \tau\partial_t y(t_{n+1/2}). \tag{7.18}$$

Hence, the result follows from Taylor's theorem. More precisely, there exist $\vartheta_1 \in (t_{n+1/2}, t_{n+1})$ and $\vartheta_2 \in (t_n, t_{n+1/2})$ such that

$$\widetilde{y}_{n+1} = y(t_{n+1/2}) + \tfrac{\tau}{2}\partial_t y(t_{n+1/2}) + \tfrac{1}{2}\tfrac{\tau^2}{4}\partial_t^2 y(t_{n+1/2}) + \tfrac{1}{6}\tfrac{\tau^3}{8}\partial_t^3 y(\vartheta_1)$$

and

$$\widetilde{y}_n = y(t_{n+1/2}) - \tfrac{\tau}{2}\partial_t y(t_{n+1/2}) + \tfrac{1}{2}\tfrac{\tau^2}{4}\partial_t^2 y(t_{n+1/2}) - \tfrac{1}{6}\tfrac{\tau^3}{8}\partial_t^3 y(\vartheta_2)$$

hold. Inserting these results in (7.8) and (7.18) yields together with $\mathrm{A} \in \mathcal{L}(\mathcal{Y}, \mathcal{X})$ the result. □

In the following subsections, we first prove wellposedness of the respective scheme together with an error estimate. Finally, this allows us to prove Theorem 7.3.

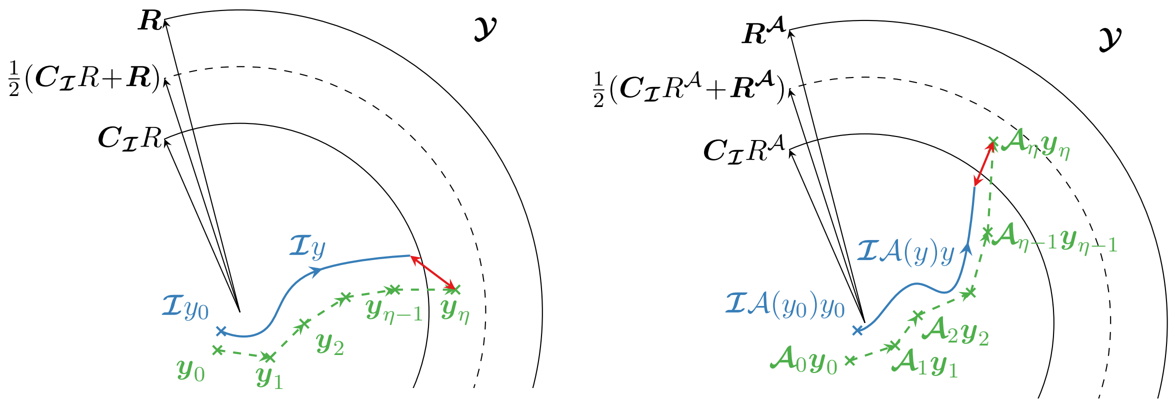

Figure 7.1: Illustration of the different radii used for the solution (left) and the differential operator applied to the solution (right).

7.1.1 Linearly implicit midpoint rule

This section is devoted to the analysis of the linearly implicit midpoint rule (7.5). In general, one would first prove the wellposedness of the scheme before tackling the error analysis. However, as in the semi-discrete setting in Section 4.3, this approach is not suitable in our case, as these proofs are intertwined here. On the one hand, we naturally need the unique existence of the next approximation to bound the error. On the other hand, we need the error estimate to prove the required bounds for the numerical solution in the $\boldsymbol{\mathcal{Y}}$-norm, before we can go to the next but one approximation. Hence, our approach is to show existence of the next approximation and the corresponding error estimate alternately by induction.

As in Section 4.3, the wellposedness of the scheme does not only imply that there exist uniquely defined iterates, but they also have to be bounded in $\boldsymbol{\mathcal{Y}}$ such that Assumption 4.1 is applicable. Therefore, we employ the same radii as in the semi-discrete case, i.e., let $R, \boldsymbol{R} > 0$ with $\boldsymbol{C_{\mathcal{I}}}R < \boldsymbol{R}$ be chosen such that both Assumptions 3.1 and 4.1 are satisfied. Let further $R^{\partial_t}, R^{\mathcal{A}}, \boldsymbol{R}^{\boldsymbol{\mathcal{A}}} > 0$ with $\boldsymbol{C_{\mathcal{I}}}R^{\mathcal{A}} < \boldsymbol{R}^{\boldsymbol{\mathcal{A}}}$ such that Assumption 7.2 is satisfied.

For $\eta < N$, the application of η steps of the linearly implicit midpoint rule is depicted in Figure 7.1. On the left-hand side, we have again the interpolation of the solution $\boldsymbol{\mathcal{I}}y$ of the continuous problem (blue), which satisfies

$$\|\boldsymbol{\mathcal{I}}y(t)\| < \boldsymbol{C_{\mathcal{I}}}R, \qquad t \in J_T,$$

due to the boundedness (4.18) of $\boldsymbol{\mathcal{I}}$ and Assumption 7.2. If the first η steps of the scheme are wellposed, this yields the existence of uniquely defined approximations $\boldsymbol{y}_1, \dots, \boldsymbol{y}_\eta \in \mathcal{Y}$ (green). Additionally, these approximations satisfy

$$\|\boldsymbol{y}_n\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{C_{\mathcal{I}}}R + \boldsymbol{R}), \qquad n = 1, \dots, \eta.$$

The difference $\boldsymbol{\mathcal{I}}y(t_\eta) - \boldsymbol{y}_\eta$ is again indicated by the red line. As in the semi-discrete case, we observe a similar behavior for $\mathcal{A}(y)y$ and $\boldsymbol{\mathcal{A}}_n\boldsymbol{y}_n$ for $n \le \eta$ with different radii. This is shown on the right-hand side.

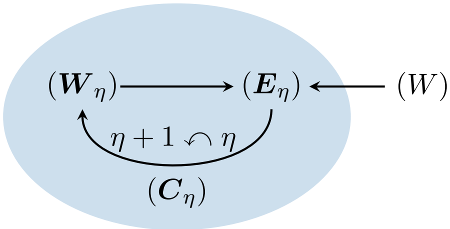

Figure 7.2: Roadmap for the analysis of the abstract fully discrete scheme with the linearly implicit midpoint rule.

Roadmap to prove wellposedness and error bounds

Our proof consists of the following steps.

- (W) From Assumption 3.2, we have wellposedness of the continuous quasilinear Cauchy problem (3.3). In particular, we get that there is a unique solution y of (3.3), which satisfies $\|y\|_{\mathcal{Y}} < R$ uniformly on J_T. The radius R is given by Assumption 3.1.

- $(\boldsymbol{W}_\eta)$ Based on the assumption that the first $\eta \geq 0$ steps are well defined, we prove in Lemma 7.7 that there exists a range of time steps $\tau \in (0, \tau^\eta_{0,\mathrm{w}})$ such that the next iterate $\boldsymbol{y}_{\eta+1}$ exists and satisfies $\|\boldsymbol{y}_{\eta+1}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}$. To be more precise, we first prove that the scheme uniquely defines $\boldsymbol{y}_{\eta+1}$ in $\boldsymbol{\mathcal{X}}$ and introduce $\tau^\eta_{0,\mathrm{w}}$ as the supremum over all time steps for which $\boldsymbol{y}_{\eta+1}$ satisfies essential bounds for the error analysis. In order to ensure $\tau^\eta_{0,\mathrm{w}} > 0$, we finally employ that $\boldsymbol{\mathcal{Y}}$ is a finite-dimensional space to provide a lower bound for $\tau^\eta_{0,\mathrm{w}}$.

- $(\boldsymbol{E}_\eta)$ In Lemma 7.8 we prove an estimate for the error $\boldsymbol{e}_{\eta+1} = \boldsymbol{\mathcal{J}} y(t_{\eta+1}) - \boldsymbol{y}_{\eta+1}$ based on the errors of the previous iterate $\boldsymbol{e}_\eta$ in $\boldsymbol{\mathcal{X}}$ using energy techniques. We further prove a corresponding result for $\boldsymbol{e}_{1/2}$ in Lemma 7.9, as the first step of the linearly implicit midpoint rule differs from the others in the approximation of the midpoint $\underline{\boldsymbol{y}}_{1/2} = \boldsymbol{y}_0$.

- $(\boldsymbol{C}_\eta)$ If for $n \leq \eta$ the numerical approximations $\boldsymbol{y}_n$ satisfy $\|\boldsymbol{y}_n - \boldsymbol{\mathcal{I}} y(t_n)\|_{\boldsymbol{\mathcal{Y}}} \to 0$ uniformly in n for $\tau, h \to 0$ under the step size restriction (7.14), we also obtain $\|\boldsymbol{y}_{\eta+1} - \boldsymbol{\mathcal{I}} y(t_{\eta+1})\|_{\boldsymbol{\mathcal{Y}}} \to 0$ using the inverse estimate (4.1).This shows that step $\eta+1$ of the linearly implicit midpoint rule applied to (4.12) is well defined. Thus, we proceed with $(\boldsymbol{W}_{\eta+1})$.

Overall, we show Theorem 7.3 by induction, as we alternately prove $(\boldsymbol{W}_\eta)$, $(\boldsymbol{E}_\eta)$, and $(\boldsymbol{C}_\eta)$. This approach is illustrated in Figure 7.2, where the analysis of the linearly implicit midpoint rule is indicated by the blue ellipse.

In the following lemma, we address the wellposedness of one step of the linearly implicit midpoint rule.

Lemma 7.7. *For $0 \leq \eta < N$ fixed assume that*

$$\|\boldsymbol{y}_\eta\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}} R), \qquad \|\boldsymbol{\mathcal{A}}_\eta \boldsymbol{y}_\eta\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}} R^{\mathcal{A}}), \qquad \|\underline{\boldsymbol{y}}_{\eta+1/2}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}. \tag{7.19}$$

If $\eta > 0$, we further assume

$$\|\underline{\boldsymbol{y}}_{\eta-1/2}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}, \qquad \|\underline{\boldsymbol{\mathcal{A}}}_{\eta-1/2}\boldsymbol{y}_{\eta-1/2}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}^{\boldsymbol{\mathcal{A}}}. \tag{7.20}$$

Then, there exists $\tau_{0,\mathrm{w}}^{\eta} > 0$, which may depend on the space discretization parameter h, such that the linearly implicit midpoint rule (7.5) has for all $\tau < \tau_{0,\mathrm{w}}^{\eta}$ a unique solution $\boldsymbol{y}_{n+1} \in \boldsymbol{\mathcal{X}}$, which satisfies

$$\|\boldsymbol{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{Y}}}, \|\boldsymbol{y}_{\eta+1}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}, \qquad \|\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2}\boldsymbol{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}^{\boldsymbol{\mathcal{A}}}.$$

Proof. The proof consists of two parts.

1. To show the existence of the next approximation $\boldsymbol{y}_{\eta+1}$ in $\boldsymbol{\mathcal{X}}$, we add $\boldsymbol{y}_\eta$ on both sides of (7.5) with $n = \eta$ and divide by 2. By (7.4), this yields

$$(\mathrm{Id} - \tfrac{\tau}{2}\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2})\boldsymbol{y}_{\eta+1/2} = \boldsymbol{y}_\eta + \tfrac{\tau}{2}\underline{\boldsymbol{\mathcal{F}}}_{\eta+1/2}. \tag{7.21}$$

   First, we observe that these expressions are well defined due to (7.19) and Assumption 4.1. Since the operator $\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2}$ is a maximal dissipative operator in $(\boldsymbol{\mathcal{X}}, \|\cdot\|_{\boldsymbol{\Lambda}(\underline{\boldsymbol{y}}_{\eta+1/2})})$, its resolvent $(\mathrm{Id} - \frac{\tau}{2}\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2})^{-1}$ exists for all $\tau > 0$, cf. [Engel and Nagel, 2000, Thm. II.3.14]. Hence, there exists a unique solution $\boldsymbol{y}_{\eta+1/2} \in \boldsymbol{\mathcal{X}}$ of (7.21), which yields $\boldsymbol{y}_{\eta+1} \in \boldsymbol{\mathcal{X}}$ by (7.4).

2. Next, we show the bounds for the approximations with respect to $\boldsymbol{\mathcal{Y}}$. Let

$$\begin{aligned}\tau_{0,\mathrm{w}}^{\eta} = \sup\{\tau_* \geq 0 \mid \|\boldsymbol{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{Y}}}, \|\boldsymbol{y}_{\eta+1}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R},\\ \|\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2}\boldsymbol{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}^{\boldsymbol{\mathcal{A}}}, \text{ for all } \tau < \tau_*\}.\end{aligned} \tag{7.22}$$

   It remains to prove $\tau_{0,\mathrm{w}}^{\eta} \geq \tau_*(h) > 0$, where $\tau_*(h)$ for $\tau_{0,\mathrm{w}}^{\eta}$ depends on the spatial discretization parameter h, but is independent of η. However, it is important to keep in mind that this lower bound only ensures $\tau_{0,\mathrm{w}}^{\eta} > 0$ for an arbitrary, but fixed space discretization parameter h. Thus, there is no necessity to keep track of the exact lower bound $\tau_*(h) > 0$, so we always take $\tau_*(h)$ as the minimum of all upper bounds for the time step we used before in order to simplify the notation, i.e., $\tau_*(h)$ is monotonically decreasing throughout the argumentation, but strictly positive.

   Since the discrete spaces are finite dimensional, we obtain from the inverse estimate (4.1), the definition (4.13) of $\boldsymbol{\mathcal{A}}$, the bound (4.16) for $\boldsymbol{\Lambda}^{-1}$, and the bound (4.29)

$$\|\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2}\|_{\mathcal{L}(\boldsymbol{\mathcal{Y}})} \leq \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}\boldsymbol{C}_{\mathbf{A}}(h)\boldsymbol{C}_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}}(h).$$

   Hence, there is $\tau_*(h) > 0$ such that the Neumann series yields

$$\begin{aligned}\|(\mathrm{Id} - \tfrac{\tau}{2}\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2})^{-1}\|_{\mathcal{L}(\boldsymbol{\mathcal{Y}})} &\leq (1 - \tfrac{\tau}{2}\|\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2}\|_{\mathcal{L}(\boldsymbol{\mathcal{Y}})})^{-1}\\ &\leq (1 - \tfrac{\tau}{2}\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}\boldsymbol{C}_{\mathbf{A}}(h)\boldsymbol{C}_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}}(h))^{-1}\\ &\leq \frac{\frac{3}{2}\boldsymbol{R} + \frac{1}{2}\boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R}{\boldsymbol{R} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R},\end{aligned}$$

for all $\tau < \tau_*(h)$, where we used $\boldsymbol{R} > \boldsymbol{C}_{\mathcal{I}} R$ in the last step. Hence, we get with (7.19)

$$\|(\mathrm{Id} - \tfrac{\tau}{2}\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2})^{-1}\boldsymbol{y}_\eta\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{4}(3\boldsymbol{R} + \boldsymbol{C}_{\mathcal{I}} R).$$

Using the boundedness (4.6) of $\mathbf{F}$ and (7.21) we deduce

$$\|\boldsymbol{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{Y}}} \le \tfrac{1}{4}(3\boldsymbol{R} + \boldsymbol{C}_{\mathcal{I}} R) + \tfrac{\tau}{2}\frac{\frac{3}{2}\boldsymbol{R} + \frac{1}{2}\boldsymbol{C}_{\mathcal{I}} R}{\boldsymbol{R} + \boldsymbol{C}_{\mathcal{I}} R}\boldsymbol{C}_{\mathbf{F}} < \boldsymbol{R},$$

for all $\tau < \tau_*(h)$, with $\tau_*(h) > 0$ sufficiently small.

To prove the bound for $\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2}\boldsymbol{y}_{\eta+1/2}$, we obtain from (7.21)

$$\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2}\boldsymbol{y}_{\eta+1/2} = (\mathrm{Id} - \tfrac{\tau}{2}\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2})^{-1}\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2}(\boldsymbol{y}_\eta + \tfrac{\tau}{2}\underline{\boldsymbol{\mathcal{F}}}_{\eta+1/2}), \tag{7.23}$$

for $\tau < \tau_*(h)$. Since $\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} > \boldsymbol{C}_{\mathcal{I}} R^{\mathcal{A}}$, we derive with similar arguments as before

$$\|(\mathrm{Id} - \tfrac{\tau}{2}\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2})^{-1}\|_{\mathcal{L}(\boldsymbol{\mathcal{Y}})} < \frac{\frac{3}{2}\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \frac{1}{2}\boldsymbol{C}_{\mathcal{I}} R^{\mathcal{A}}}{\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \boldsymbol{C}_{\mathcal{I}} R^{\mathcal{A}}}, \tag{7.24}$$

for $\tau < \tau_*(h)$. Using the triangle inequality and the inverse estimate (4.1), we compute

$$\|\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2}\boldsymbol{y}_\eta\|_{\boldsymbol{\mathcal{Y}}} \le \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\|(\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2} - \boldsymbol{\mathcal{A}}_\eta)\boldsymbol{y}_\eta\|_{\boldsymbol{\mathcal{X}}} + \|\boldsymbol{\mathcal{A}}_\eta\boldsymbol{y}_\eta\|_{\boldsymbol{\mathcal{Y}}},$$

where we further employ (4.14) to get

$$\begin{aligned}\|\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2}\boldsymbol{y}_\eta\|_{\boldsymbol{\mathcal{Y}}} &\le \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\boldsymbol{L}_{\boldsymbol{\mathcal{A}}}\|\boldsymbol{\mathcal{A}}_\eta\boldsymbol{y}_\eta\|_{\boldsymbol{\mathcal{Y}}}\|\underline{\boldsymbol{y}}_{\eta+1/2} - \boldsymbol{y}_\eta\|_{\boldsymbol{\mathcal{X}}} + \|\boldsymbol{\mathcal{A}}_\eta\boldsymbol{y}_\eta\|_{\boldsymbol{\mathcal{Y}}} \\ &\le (1 + \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\boldsymbol{L}_{\boldsymbol{\mathcal{A}}}\|\underline{\boldsymbol{y}}_{\eta+1/2} - \boldsymbol{y}_\eta\|_{\boldsymbol{\mathcal{X}}})\|\boldsymbol{\mathcal{A}}_\eta\boldsymbol{y}_\eta\|_{\boldsymbol{\mathcal{Y}}}.\end{aligned}$$

We now have to consider the following two cases depending on η. If $\eta = 0$, we have $\underline{\boldsymbol{y}}_{1/2} = \boldsymbol{y}_0$ and hence (7.19) yields $\|\underline{\boldsymbol{\mathcal{A}}}_{1/2}\boldsymbol{y}_0\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}^{\boldsymbol{\mathcal{A}}}$. For $\eta > 0$, we further compute with (7.20) and (4.6)

$$\begin{aligned}\|\underline{\boldsymbol{y}}_{\eta+1/2} - \boldsymbol{y}_\eta\|_{\boldsymbol{\mathcal{X}}} &= \tfrac{1}{2}\|\boldsymbol{y}_\eta - \boldsymbol{y}_{\eta-1}\|_{\boldsymbol{\mathcal{X}}} \\ &= \tfrac{\tau}{2}\|\underline{\boldsymbol{\mathcal{A}}}_{\eta-1/2}\boldsymbol{y}_{\eta-1/2} + \underline{\boldsymbol{\mathcal{F}}}_{\eta-1/2}\|_{\boldsymbol{\mathcal{X}}} \\ &\le \tfrac{\tau}{2}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \boldsymbol{C}_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}}(h)\boldsymbol{C}_{\mathbf{F}}).\end{aligned}$$

Thus, we have shown the bound

$$\|\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2}\boldsymbol{y}_\eta\|_{\boldsymbol{\mathcal{Y}}} \le (1 + \tfrac{\tau}{2}C_{\boldsymbol{\mathcal{A}}\boldsymbol{y}}(h))\|\boldsymbol{\mathcal{A}}_\eta\boldsymbol{y}_\eta\|_{\boldsymbol{\mathcal{Y}}}, \tag{7.25}$$

with a constant $C_{\boldsymbol{\mathcal{A}}\boldsymbol{y}}(h) = \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\boldsymbol{L}_{\boldsymbol{\mathcal{A}}}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \boldsymbol{C}_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}}(h)\boldsymbol{C}_{\mathbf{F}})$.

Finally, we obtain from (4.1), the definition (4.13) of $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{F}}$, as well as the bounds (4.6), (4.16), and (4.29) for $\mathbf{F}$, $\boldsymbol{\Lambda}^{-1}$, and $\mathbf{A}$, respectively, the estimate

$$\|\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2}\underline{\boldsymbol{\mathcal{F}}}_{\eta+1/2}\|_{\boldsymbol{\mathcal{Y}}} \le C_{\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{F}}}(h), \tag{7.26}$$

where the constant is given by $C_{\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{F}}}(h) = \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}\boldsymbol{C}_{\mathbf{A}}(h)\boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}\boldsymbol{C}_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}}(h)\boldsymbol{C}_{\mathbf{F}}$.

Hence, using the bounds (7.19), (7.24), (7.25), and (7.26) in (7.23) proves

$$\|\underline{\mathcal{A}}_{\eta+1/2}\boldsymbol{y}_{\eta+1/2}\|_{\mathcal{Y}} < (1+\tfrac{\tau}{2}C_{\mathcal{A}\boldsymbol{y}}(h))\tfrac{1}{4}(3\boldsymbol{R}^{\mathcal{A}} + \boldsymbol{C}_{\mathcal{I}}R^{\mathcal{A}}) + \tfrac{\tau}{2}\frac{\tfrac{3}{2}\boldsymbol{R}^{\mathcal{A}} + \tfrac{1}{2}\boldsymbol{C}_{\mathcal{I}}R^{\mathcal{A}}}{\boldsymbol{R}^{\mathcal{A}} + \boldsymbol{C}_{\mathcal{I}}R^{\mathcal{A}}}C_{\mathcal{AF}}(h).$$

As before, this yields the existence of $\tau_*(h) > 0$ such that

$$\|\underline{\mathcal{A}}_{\eta+1/2}\boldsymbol{y}_{\eta+1/2}\|_{\mathcal{Y}} < \boldsymbol{R}^{\mathcal{A}} \tag{7.27}$$

holds for all $\tau < \tau_*(h)$.

Combination of (7.19) and (7.27) with (4.6) in (7.5) finally yields the bound for $\|\boldsymbol{y}_{n+1}\|_{\mathcal{Y}}$ for all $\tau < \tau_*(h)$ with some $\tau_*(h) > 0$.

Thus, we have shown $\tau_{0,\mathrm{w}}^{\eta} \geq \tau_*(h) > 0$.

Finally, we point out that the estimates in the second part of this proof are far from sharp, as we see at the end of this section. Nevertheless, they are necessary and sufficient to ensure the wellposedness of the scheme. □

As we have seen in the previous lemma, there is an interval $(0, \tau_{0,\mathrm{w}}^{\eta})$ such that all time-step sizes taken from this interval yield a wellposed scheme (7.5). This allows us to tackle the error of the scheme in the next lemma, i.e., we bound the error after one step of the linearly implicit midpoint rule with respect to the errors at previous time steps and discretization errors in space. We keep track of the constants appearing with the linearization $\underline{\boldsymbol{y}}_{\eta+1/2}$, as they are important in the next section, where we analyze the fully implicit midpoint rule.

Lemma 7.8. *Let Assumption 7.2 be true. If the assumptions of Lemma 7.7 are satisfied for $0 \leq \eta < N$ fixed, the error of the linearly implicit midpoint rule satisfies for $\tau < \tau_{0,\mathrm{w}}^{\eta}$ the bound*

$$\begin{aligned}
\|\boldsymbol{e}_{\eta+1}\|^2_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1})} &\leq (1+C\tau)\|\boldsymbol{e}_\eta\|^2_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_\eta)} + C_{\underline{\boldsymbol{e}}}\tau\|\mathcal{J}\widetilde{y}_{\eta+1/2} - \underline{\boldsymbol{y}}_{\eta+1/2}\|^2_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1/2})} \\
&\quad + C\tau^5\Big(\sup_{[t_\eta,t_{\eta+1}]}\|\partial_t^3 y\|^2_{\mathcal{X}} + \sup_{[t_{\eta-1},t_{\eta+1}]}\|\partial_t^2 y\|^2_{\mathcal{Y}}\Big) + C\tau\|(\mathcal{I}-\mathcal{J})\widetilde{y}_{\eta+1/2}\|^2_{\mathcal{X}} \\
&\quad + C\tau\Big(\sup_{s_1,s_2\in[t_\eta,t_{\eta+1}]}\|\mathcal{R}_\Lambda(y(s_1))\partial_t y(s_2)\|^2_{\mathcal{X}} + \sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_{\mathrm{A}} y\|^2_{\mathcal{X}} + \sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_{\mathrm{F}}(\cdot,y)\|^2_{\mathcal{X}}\Big),
\end{aligned} \tag{7.28}$$

with constants $C, C_{\underline{\boldsymbol{e}}} > 0$ independent of η, h and τ.

Proof. For the sake of presentation, we use the notation

$$(\cdot \mid \cdot)_* = (\cdot \mid \cdot)_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1/2})}, \qquad \|\cdot\|_* = \|\cdot\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1/2})}$$

throughout this proof. Following the approach presented in [Hochbruck et al., 2018, Lem. 5.1], the proof is based on energy techniques. First, using (7.9) together with (7.10) yields

$$\begin{aligned}
\|\boldsymbol{e}_{\eta+1}\|^2_* - \|\boldsymbol{e}_\eta\|^2_* &= \Big(\boldsymbol{e}_{\eta+1} + \boldsymbol{e}_\eta \mid \boldsymbol{e}_{\eta+1} - \boldsymbol{e}_\eta\Big)_* \\
&= 2\tau\Big(\boldsymbol{e}_{\eta+1/2} \mid \widetilde{\mathcal{A}}_{\eta+1/2}\boldsymbol{e}_{\eta+1/2} + \boldsymbol{g}^{\mathrm{LI}}_{\eta+1}\Big)_*.
\end{aligned} \tag{7.29}$$

Due to the definition (4.13) of $\boldsymbol{\mathcal{A}}$ and the dissipativity (4.5) of $\mathbf{A}$, we obtain

$$\Big(\boldsymbol{e}_{\eta+1/2} \mid \widetilde{\mathcal{A}}_{\eta+1/2}\boldsymbol{e}_{\eta+1/2}\Big)_* = \Big(\boldsymbol{e}_{\eta+1/2} \mid \mathbf{A}\boldsymbol{e}_{\eta+1/2}\Big)_{\mathcal{X}} \leq 0.$$

Thus, we get with Young's inequality

$$\|\boldsymbol{e}_{\eta+1}\|_*^2 - \|\boldsymbol{e}_\eta\|_*^2 \le \tau\|\boldsymbol{e}_{\eta+1/2}\|_*^2 + \tau\|\boldsymbol{g}_{\eta+1}^{\mathrm{LI}}\|_*^2. \tag{7.30}$$

To further bound these terms, we use again (7.10), (4.13), (4.5), and the Cauchy–Schwarz inequality to compute

$$\begin{aligned}\|\boldsymbol{e}_{\eta+1/2}\|_*^2 &= \Big(\boldsymbol{e}_{\eta+1/2} \mid \boldsymbol{e}_\eta\Big)_* + \tfrac{\tau}{2}\Big(\boldsymbol{e}_{\eta+1/2} \mid \widetilde{\boldsymbol{\mathcal{A}}}_{\eta+1/2}\boldsymbol{e}_{\eta+1/2} + \boldsymbol{g}_{\eta+1}^{\mathrm{LI}}\Big)_* \\ &\le \|\boldsymbol{e}_{\eta+1/2}\|_*\Big(\|\boldsymbol{e}_\eta\|_* + \tfrac{\tau}{2}\|\boldsymbol{g}_{\eta+1}^{\mathrm{LI}}\|_*\Big).\end{aligned}$$

Hence, we have

$$\|\boldsymbol{e}_{\eta+1/2}\|_* \le \|\boldsymbol{e}_\eta\|_* + \tfrac{\tau}{2}\|\boldsymbol{g}_{\eta+1}^{\mathrm{LI}}\|_*, \tag{7.31}$$

which, together with (7.30), implies

$$\|\boldsymbol{e}_{\eta+1}\|_*^2 \le (1 + C\tau)\|\boldsymbol{e}_\eta\|_*^2 + C\tau\|\boldsymbol{g}_{\eta+1}^{\mathrm{LI}}\|_*^2 \tag{7.32}$$

with a constant $C > 0$ independent of τ and h. We now focus on $\|\boldsymbol{g}_{\eta+1}^{\mathrm{LI}}\|_*$. First, the definition (7.12) of $\boldsymbol{g}_{\eta+1}^{\mathrm{LI}}$ yields with the triangle inequality

$$\begin{aligned}\|\boldsymbol{g}_{\eta+1}^{\mathrm{LI}}\|_* \le{}& \|(\widetilde{\boldsymbol{\mathcal{A}}}_{\eta+1/2} - \underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2})\boldsymbol{y}_{\eta+1/2}\|_* + \|\widetilde{\boldsymbol{\mathcal{F}}}_{\eta+1/2} - \underline{\boldsymbol{\mathcal{F}}}_{\eta+1/2}\|_* \\ &+ \|(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{\eta+1/2}])\tfrac{1}{\tau}(\widetilde{y}_{\eta+1} - \widetilde{y}_\eta)\|_* + \|\boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{\eta+1/2}]\widetilde{\mathcal{F}}_{\eta+1/2} - \widetilde{\boldsymbol{\mathcal{F}}}_{\eta+1/2}\|_* \\ &+ \|(\widetilde{\boldsymbol{\mathcal{A}}}_{\eta+1/2}\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{\eta+1/2}]\widetilde{\mathcal{A}}_{\eta+1/2})(\widehat{\delta}_{\eta+1/2} - \widetilde{y}_{\eta+1/2})\|_* \\ &+ \|\boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{\eta+1/2}]\widetilde{\mathcal{A}}_{\eta+1/2})\widehat{\delta}_{\eta+1/2}\|_* + \|\tfrac{1}{\tau}\boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{\eta+1/2}]\delta_{\eta+1}\|_*.\end{aligned} \tag{7.33}$$

We consider all terms separately. For the first term, we have with (4.10), (4.14), and Lemma 7.7 for $\tau < \tau_{0,\mathrm{w}}^\eta$ the bound

$$\begin{aligned}\|(\widetilde{\boldsymbol{\mathcal{A}}}_{\eta+1/2} - \underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2})\boldsymbol{y}_{\eta+1/2}\|_* &\le \boldsymbol{C}_{\boldsymbol{\Lambda}}^{\frac{1}{2}}\boldsymbol{L}_{\boldsymbol{\mathcal{A}}}\|\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2}\boldsymbol{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{Y}}}\|\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2} - \underline{\boldsymbol{y}}_{\eta+1/2}\|_{\boldsymbol{\mathcal{X}}} \\ &\le \boldsymbol{C}_{\boldsymbol{\Lambda}}^{\frac{1}{2}}\boldsymbol{L}_{\boldsymbol{\mathcal{A}}}\boldsymbol{R}^{\boldsymbol{\mathcal{A}}}(\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\widetilde{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{X}}} + \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-\frac{1}{2}}\|\boldsymbol{\mathcal{J}}\widetilde{y}_{\eta+1/2} - \underline{\boldsymbol{y}}_{\eta+1/2}\|_*).\end{aligned}$$

Similarly, we derive for the second term with (4.10) and (4.15)

$$\begin{aligned}\|\widetilde{\boldsymbol{\mathcal{F}}}_{\eta+1/2} - \underline{\boldsymbol{\mathcal{F}}}_{\eta+1/2}\|_* &\le \boldsymbol{C}_{\boldsymbol{\Lambda}}^{\frac{1}{2}}\boldsymbol{L}_{\boldsymbol{\mathcal{F}}}\|\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2} - \underline{\boldsymbol{y}}_{\eta+1/2}\|_{\boldsymbol{\mathcal{X}}} \\ &\le \boldsymbol{C}_{\boldsymbol{\Lambda}}^{\frac{1}{2}}\boldsymbol{L}_{\boldsymbol{\mathcal{F}}}(\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\widetilde{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{X}}} + \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-\frac{1}{2}}\|\boldsymbol{\mathcal{J}}\widetilde{y}_{\eta+1/2} - \underline{\boldsymbol{y}}_{\eta+1/2}\|_*).\end{aligned}$$

Due to (4.23), we further get for the last terms

$$\|\boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{\eta+1/2}]\widetilde{\mathcal{A}}_{\eta+1/2})\widehat{\delta}_{\eta+1/2}\|_* + \|\tfrac{1}{\tau}\boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{\eta+1/2}]\delta_{\eta+1}\|_* \le C(\|\mathrm{A}\widehat{\delta}_{\eta+1/2}\|_{\mathcal{X}} + \|\tfrac{1}{\tau}\delta_{\eta+1}\|_{\mathcal{X}}).$$

As the other terms in (7.33) are covered by Lemma 4.19, we get with

$$\widetilde{y}_{\eta+1} - \widetilde{y}_\eta = \int_{t_\eta}^{t_{\eta+1}} \partial_t y(s)\,\mathrm{d}s$$

and (7.8) for the right-hand side

$$\begin{aligned}\|\boldsymbol{g}_{\eta+1}^{\mathrm{LI}}\|_* \le C\big(&\|(\boldsymbol{\mathcal{I}}-\boldsymbol{\mathcal{J}})\widetilde{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{X}}} + \|\boldsymbol{\mathcal{J}}\widetilde{y}_{\eta+1/2}-\underline{\boldsymbol{y}}_{\eta+1/2}\|_* + \sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_\Lambda(\widetilde{y}_{\eta+1/2})\partial_t y\|_{\boldsymbol{\mathcal{X}}}\\ &+\|\mathcal{R}_\Lambda(\widetilde{y}_{\eta+1}+\widetilde{y}_\eta)\|_{\boldsymbol{\mathcal{X}}} + \|\mathcal{R}_\mathrm{F}(t_{\eta+1/2},\widetilde{y}_{\eta+1/2})\|_{\boldsymbol{\mathcal{X}}} + \|\mathrm{A}\widehat{\delta}_{\eta+1/2}\|_{\mathcal{X}} + \|\tfrac{1}{\tau}\delta_{\eta+1}\|_{\mathcal{X}}\big).\end{aligned} \tag{7.34}$$

Finally, the norm equivalence (4.11) yields

$$\begin{aligned}\|\boldsymbol{e}_{\eta+1}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1})} &\le (1+C\tau)\|\boldsymbol{e}_{\eta+1}\|_*^2,\\ \|\boldsymbol{e}_\eta\|_*^2 &\le (1+C\tau)\|\boldsymbol{e}_\eta\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_\eta)}.\end{aligned} \tag{7.35}$$

Hence, (7.32) and (7.34) imply

$$\begin{aligned}\|\boldsymbol{e}_{\eta+1}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1})} \le (1+C\tau)\|\boldsymbol{e}_\eta\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_\eta)} + C\tau\Big(&\|(\boldsymbol{\mathcal{I}}-\boldsymbol{\mathcal{J}})\widetilde{y}_{\eta+1/2}\|^2_{\boldsymbol{\mathcal{X}}} + \|\boldsymbol{\mathcal{J}}\widetilde{y}_{\eta+1/2}-\underline{\boldsymbol{y}}_{\eta+1/2}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})}\\ &+\sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_\Lambda(\widetilde{y}_{\eta+1/2})\partial_t y\|^2_{\boldsymbol{\mathcal{X}}} + \|\mathcal{R}_\Lambda(\widetilde{y}_{\eta+1}+\widetilde{y}_\eta)\|^2_{\boldsymbol{\mathcal{X}}} + \|\mathcal{R}_\mathrm{F}(t_{\eta+1/2},\widetilde{y}_{\eta+1/2})\|^2_{\boldsymbol{\mathcal{X}}}\\ &+\|\mathrm{A}\widehat{\delta}_{\eta+1/2}\|^2_{\mathcal{X}} + \|\tfrac{1}{\tau}\delta_{\eta+1}\|^2_{\mathcal{X}}\Big).\end{aligned}$$

Due to $\mathrm{A}\in\mathcal{L}(\mathcal{Y},\mathcal{X})$, the estimates for the defects from Lemma 7.6 conclude the proof. □

Using the preliminary lemmas, we are now able to provide an error estimate for the linearly implicit midpoint rule. However, we have to take special care of the first step, since we derive from the Taylor's theorem and the continuity (4.11) of the state-dependent norm

$$\begin{aligned}\tau\|\boldsymbol{\mathcal{J}}\widetilde{y}_{1/2}-\underline{\boldsymbol{y}}_{1/2}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{1/2})} &\le \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}\boldsymbol{C}_{\boldsymbol{\mathcal{J}}}^2\tau\|\widetilde{y}_{1/2}-\widetilde{y}_0\|^2_{\mathcal{Y}} + \tau\|\boldsymbol{\mathcal{J}}\widetilde{y}_0-\boldsymbol{y}_0\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{1/2})}\\ &\le \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}\boldsymbol{C}_{\boldsymbol{\mathcal{J}}}^2\tfrac{\tau^3}{4}\sup_{[t_0,t_{1/2}]}\|\partial_t y\|^2_{\mathcal{Y}} + (1+C'\tfrac{\tau}{2})^2\tau\|\boldsymbol{e}_0\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_0)}.\end{aligned} \tag{7.36}$$

Nevertheless, we would still get convergence of order $\frac{3}{2}$ in time for the global error, as we do not need Gronwall's inequality for the first step. However, as the rest of the scheme is of order 2 in time, this is still not satisfactory.

Hence, we next provide an alternative error estimate, which is used later only for the first step. In particular, compared to (7.28) we allow for a larger factor to be multiplied by $\|\boldsymbol{e}_\eta\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_\eta)}$ in order to gain powers of τ for the other terms.

Lemma 7.9. *Let Assumption 7.2 be true. If the assumptions of Lemma 7.7 are satisfied for $0\le\eta<N$ fixed, the error of the linearly implicit midpoint rule satisfies for $\tau<\tau_{0,\mathrm{w}}^\eta$ the bound*

$$\begin{aligned}\|\boldsymbol{e}_{\eta+1}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1})} \le C\|\boldsymbol{e}_\eta\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_\eta)} &+ C_{\underline{e}}\tau^2\|\boldsymbol{\mathcal{J}}\widetilde{y}_{\eta+1/2}-\underline{\boldsymbol{y}}_{\eta+1/2}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})}\\ &+C\tau^6\Big(\sup_{[t_\eta,t_{\eta+1}]}\|\partial_t^3 y\|^2_{\mathcal{X}} + \sup_{[t_{\eta-1},t_{\eta+1}]}\|\partial_t^2 y\|^2_{\mathcal{Y}}\Big) + C\tau^2\|(\boldsymbol{\mathcal{I}}-\boldsymbol{\mathcal{J}})\widetilde{y}_{\eta+1/2}\|^2_{\boldsymbol{\mathcal{X}}}\\ &+C\tau^2\Big(\sup_{s_1,s_2\in[t_\eta,t_{\eta+1}]}\|\mathcal{R}_\Lambda(y(s_1))\partial_t y(s_2)\|^2_{\boldsymbol{\mathcal{X}}} + \sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_\Lambda y\|^2_{\boldsymbol{\mathcal{X}}} + \sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_\mathrm{F}(\cdot,y)\|^2_{\boldsymbol{\mathcal{X}}}\Big),\end{aligned} \tag{7.37}$$

with constants $C, C_{\underline{e}}>0$ independent of h and τ.

Proof. To prove this result, we proceed mostly as in the proof of Lemma 7.8, but modify the treatment of the time-step size τ. From (7.29), using different weights in Young's inequality, we get

$$\|\boldsymbol{e}_{\eta+1}\|_*^2 - \|\boldsymbol{e}_\eta\|_*^2 \leq \|\boldsymbol{e}_{\eta+1/2}\|_*^2 + \tau^2 \|\boldsymbol{g}_{\eta+1}^{\mathrm{LI}}\|_*^2.$$

instead of (7.30) and thus with (7.31)

$$\|\boldsymbol{e}_{\eta+1}\|_*^2 \leq C\|\boldsymbol{e}_\eta\|_*^2 + C\tau^2 \|\boldsymbol{g}_{\eta+1}^{\mathrm{LI}}\|_*^2,$$

instead of (7.32). Using now the same arguments as in the proof of Lemma 7.8 yields the result. □

Based on these lemmas, we now prove the main result of this section.

Proof of Theorem 7.3. The proof is done by induction, where we alternately use Lemma 7.7 to prove existence of the next approximations and Lemma 7.8 to prove the error bound

$$\begin{aligned}
\|\boldsymbol{e}_n\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_n)}^2 &\leq C\|\boldsymbol{e}_0\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_0)}^2 + C\tau\sum_{r=0}^{n-1}\|\boldsymbol{e}_r\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_r)}^2 + Ct_n \sup_{[0,t_n]}\|(\boldsymbol{\mathcal{I}}-\boldsymbol{\mathcal{J}})y\|_{\boldsymbol{\mathcal{X}}}^2 \\
&\quad + Ct_n\tau^4\Big(\sup_{[0,t_n]}\|\partial_t y\|_{\mathcal{Y}}^2 + \sup_{[0,t_n]}\|\partial_t^2 y\|_{\mathcal{Y}}^2 + \sup_{[0,t_n]}\|\partial_t^3 y\|_{\mathcal{X}}^2\Big) \qquad (7.38)\\
&\quad + Ct_n\Big(\sup_{s_1,s_2\in[0,t_n]}\|\mathcal{R}_\Lambda(y(s_1))\partial_t y(s_2)\|_{\boldsymbol{\mathcal{X}}}^2 + \sup_{[0,t_n]}\|\mathcal{R}_{\mathrm{A}} y\|_{\boldsymbol{\mathcal{X}}}^2 + \sup_{[0,t_n]}\|\mathcal{R}_{\mathrm{F}}(\cdot,y)\|_{\boldsymbol{\mathcal{X}}}^2\Big)
\end{aligned}$$

for $n = 0,\dots,N$. To conclude, we show that, based on the error estimate, we can improve the upper bound for the norm of the next iterates in $\boldsymbol{\mathcal{Y}}$ such that they again satisfy the assumptions of Lemma 7.7, and go to the next step.

For $n = 0$, Assumption 4.22 yields the existence of $h_1 > 0$ such that the initial value satisfies

$$\|\boldsymbol{y}_0\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R), \qquad \|\boldsymbol{\mathcal{A}}(\boldsymbol{y}_0)\boldsymbol{y}_0\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R^{\mathcal{A}}),$$

for all $h < h_1$. Thus, Lemma 7.7 is applicable, since $\underline{\boldsymbol{y}}_{1/2} = \boldsymbol{y}_0 \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$. Then, (7.37) together with (7.36) yields the result.

For the induction step, we assume that the assumptions of Lemma 7.7 are satisfied and the error bound (7.38) holds true up to some $n = \eta \in \{0,\dots,N-1\}$ arbitrary but fixed. Hence, the assumptions of Lemma 7.8 are also satisfied. To close the induction argument, we show that this is also the case for $n = \eta + 1$.

First, Lemma 7.7 yields the existence of

$$\boldsymbol{y}_{\eta+1/2}, \boldsymbol{y}_{\eta+1} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}), \qquad \underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2}\boldsymbol{y}_{\eta+1/2} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}}).$$

In order to use the estimate from Lemma 7.8, we derive from (7.6) and the triangle inequality

$$\begin{aligned}
\|\boldsymbol{\mathcal{J}}\widetilde{y}_{\eta+1/2} - \underline{\boldsymbol{y}}_{\eta+1/2}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} &\leq \tfrac{3}{2}\|\boldsymbol{\mathcal{J}}\widetilde{y}_\eta - \boldsymbol{y}_\eta\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} + \tfrac{1}{2}\|\boldsymbol{\mathcal{J}}\widetilde{y}_{\eta-1} - \boldsymbol{y}_{\eta-1}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} \\
&\quad + \tfrac{1}{2}\|\boldsymbol{\mathcal{J}}(2\widetilde{y}_{\eta+1/2} - 3\widetilde{y}_\eta + \widetilde{y}_{\eta-1})\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} \\
&\leq \tfrac{3}{2}\|\boldsymbol{e}_\eta\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} + \tfrac{1}{2}\|\boldsymbol{e}_{\eta-1}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} + C\tau^2 \sup_{[t_{\eta-1},t_{\eta+1/2}]}\|\partial_t^2 y\|_{\mathcal{Y}},
\end{aligned} \qquad (7.39)$$

where we used (4.10), (4.17), and Taylor's theorem in the last step. Due to (4.11), we get

$$\|\boldsymbol{e}_{\eta-1}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} \leq (1 + C\tau)\|\boldsymbol{e}_{\eta-1}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta-1})}.$$

Hence, this yields together with the norm equivalences (7.35) and the error estimate (7.28) the bound

$$\begin{aligned}\|\boldsymbol{e}_{\eta+1}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1})} &\leq \|\boldsymbol{e}_{\eta}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta})} + C\tau\|\boldsymbol{e}_{\eta}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta})} + C\tau\|\boldsymbol{e}_{\eta-1}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta-1})} \\ &\quad + C\tau^5\Big(\sup_{[t_\eta,t_{\eta+1}]}\|\partial_t^3 y\|^2_{\mathcal{X}} + \sup_{[t_{\eta-1},t_{\eta+1}]}\|\partial_t^2 y\|^2_{\mathcal{Y}}\Big) + C\tau\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\widetilde{y}_{\eta+1/2}\|^2_{\boldsymbol{\mathcal{X}}} \\ &\quad + C\tau\Big(\sup_{s_1,s_2\in[t_\eta,t_{\eta+1}]}\|\mathcal{R}_\Lambda(y(s_1))\partial_t y(s_2)\|^2_{\boldsymbol{\mathcal{X}}} + \sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_{\mathrm{A}} y\|^2_{\boldsymbol{\mathcal{X}}} + \sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_{\mathrm{F}}(\cdot,y)\|^2_{\boldsymbol{\mathcal{X}}}\Big).\end{aligned}$$

Using the induction hypothesis to replace $\|\boldsymbol{e}_{\eta}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta})}$, we further get

$$\begin{aligned}\|\boldsymbol{e}_{\eta+1}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1})} &\leq C\|\boldsymbol{e}_{0}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{0})} + C\tau\sum_{r=0}^{\eta}\|\boldsymbol{e}_{r}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{r})} \\ &\quad + Ct_{\eta+1}\tau^4\Big(\sup_{[0,t_{\eta+1}]}\|\partial_t^3 y\|^2_{\mathcal{X}} + \sup_{[0,t_{\eta+1}]}\|\partial_t^2 y\|^2_{\mathcal{Y}}\Big) + Ct_{\eta+1}\sup_{[0,t_{\eta+1}]}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})y\|^2_{\boldsymbol{\mathcal{X}}} \\ &\quad + Ct_{\eta+1}\Big(\sup_{s_1,s_2\in[0,t_{\eta+1}]}\|\mathcal{R}_\Lambda(y(s_1))\partial_t y(s_2)\|^2_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_{\eta+1}]}\|\mathcal{R}_{\mathrm{A}} y\|^2_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_{\eta+1}]}\|\mathcal{R}_{\mathrm{F}}(\cdot,y)\|^2_{\boldsymbol{\mathcal{X}}}\Big).\end{aligned}$$

Finally, the discrete Gronwall inequality and taking the square root yields the error estimate (7.15) for $n = \eta + 1$.

To conclude the proof, we have to ensure that the assumptions of Lemma 7.7 are also satisfied for the next step $n = \eta + 1$, i.e., we have to ensure

$$\|\boldsymbol{y}_{\eta+1}\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R), \qquad \|\boldsymbol{\mathcal{A}}_{\eta+1}\boldsymbol{y}_{\eta+1}\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R^{\mathcal{A}}), \tag{7.40}$$

and

$$\|\underline{\boldsymbol{y}}_{\eta+3/2}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}. \tag{7.41}$$

Furthermore, we have to provide $\tau_0, h_0 > 0$ independent of η such that the linearly implicit midpoint rule is wellposed and (7.15) is satisfied for all $\tau < \tau_0$ and $h < h_0$ under the step size restriction (7.14). In particular, this then implies $\tau_{0,\mathrm{w}}^{\eta} \geq \tau_0$.

For (7.40), we get from the error estimate (7.15) and Assumption 4.22

$$C_{\max}(h)\|\boldsymbol{e}_{\eta+1}\|_{\boldsymbol{\mathcal{X}}} \to 0,$$

for $\tau, h \to 0$ satisfying the step size restriction (7.14). Thus, Lemma 4.24 directly yields the existence of $\tau_2, h_2 > 0$ such that (7.40) is satisfied for all $\tau < \tau_2$ and $h < h_2$ satisfying the step size restriction (7.14).

To prove (7.41), we proceed as in (7.39) to compute

$$\|\boldsymbol{\mathcal{J}}\widetilde{y}_{\eta+3/2} - \underline{\boldsymbol{y}}_{\eta+3/2}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+3/2})} \leq C\|\boldsymbol{e}_{\eta+1}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+3/2})} + C\|\boldsymbol{e}_{\eta}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+3/2})} + C\tau^2\sup_{[t_\eta,t_{\eta+3/2}]}\|\partial_t^2 y\|_{\mathcal{Y}}.$$

Thus, the inverse estimate (4.1) and the norm equivalence (4.10) yield

$$\|\mathcal{J}\widetilde{y}_{\eta+3/2} - \underline{\boldsymbol{y}}_{\eta+3/2}\|_{\boldsymbol{\mathcal{Y}}} \leq C\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\Big(\|\boldsymbol{e}_{\eta+1}\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1})} + \|\boldsymbol{e}_{\eta}\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta})} + \tau^2 \sup_{[t_\eta,t_{\eta+3/2}]} \|\partial_t^2 y\|_{\mathcal{Y}}\Big).$$

Note that the error estimate (7.15), which we just proved to be true for $n = \eta+1$, together with (4.1) and Assumption 4.22 yield that there exist $\tau_3, h_3 > 0$ independent of η such that

$$\begin{aligned}\|\mathcal{I}\widetilde{y}_{\eta+3/2} - \underline{\boldsymbol{y}}_{\eta+3/2}\|_{\boldsymbol{\mathcal{Y}}} &\leq \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\|(\mathcal{I}-\mathcal{J})\widetilde{y}_{\eta+3/2}\|_{\boldsymbol{\mathcal{X}}} + \|\mathcal{J}\widetilde{y}_{\eta+3/2} - \underline{\boldsymbol{y}}_{\eta+3/2}\|_{\boldsymbol{\mathcal{Y}}} \\ &\leq \boldsymbol{R} - \boldsymbol{C}_{\mathcal{I}}R\end{aligned}$$

holds for all $\tau < \tau_3$ and $h < h_3$ satisfying the step size restriction (7.14), as all terms in the first line vanish for $\tau, h \to 0$. This yields (7.41), since we have

$$\|\underline{\boldsymbol{y}}_{\eta+3/2}\|_{\boldsymbol{\mathcal{Y}}} \leq \|\mathcal{I}\widetilde{y}_{\eta+3/2}\|_{\boldsymbol{\mathcal{Y}}} + \|\mathcal{I}\widetilde{y}_{\eta+3/2} - \underline{\boldsymbol{y}}_{\eta+3/2}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}.$$

Due to (7.39), Lemma 4.24 further yields the existence of $\tau_4, h_4 > 0$ such that

$$\|\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2}\boldsymbol{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}^{\boldsymbol{\mathcal{A}}}$$

holds for all $\tau < \tau_4$ and $h < h_4$ satisfying the step size restriction (7.14). Thus, as (7.40) and (7.19) yield

$$\|\boldsymbol{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{Y}}}, \|\boldsymbol{y}_{\eta+1}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R},$$

for $\tau < \min\{\tau_2, \tau_3\}$ and $h < \min\{h_2, h_3\}$ under the step size restriction (7.14), we finally define

$$\tau_0 := \min\{\tau_2, \tau_3, \tau_4\}, \qquad h_0 := \min\{h_1, h_2, h_3, h_4\},$$

which concludes the proof. □

In the next subsection, we use the results shown for the linearly implicit midpoint rule to analyze also the fully implicit midpoint rule.

7.1.2 Fully implicit midpoint rule

We now focus on the fully implicit midpoint rule (7.3). Here, the main difficulty is to prove the existence of the next iterates, even under the step size restriction (7.14). For the linearly implicit midpoint rule, the corresponding wellposedness result is shown in Lemma 7.7. However, since we take the supremum in (7.22) over all time steps for which the scheme is wellposed, the proof thereof crucially depends on the linearity of the scheme. Hence, this is not directly applicable for the implicit treatment of the nonlinearity, as the existence of the iterates for nonlinear schemes is not even guaranteed for slight variations of the time step. In particular, if we follow the same approach as in the previous section, we would be stuck with the very restrictive step size restriction used in the proof of Lemma 7.7 to ensure $\tau_{0,\mathrm{w}} > 0$.

To circumvent these difficulties, we use a fixed-point iteration in every time step to derive a sequence of linear problems. These are covered by the results from the previous section. We consider for $\eta < N$ fixed the sequence $(\boldsymbol{y}_{\eta+1}^k)_{k\in\mathbb{N}_0}$, which is recursively defined by

$$\boldsymbol{y}_{\eta+1}^{k+1} = \boldsymbol{y}_\eta + \tau\boldsymbol{\mathcal{A}}_{\eta+1/2}^k\boldsymbol{y}_{\eta+1/2}^{k+1} + \tau\boldsymbol{\mathcal{F}}_{\eta+1/2}^k, \qquad k \geq 0, \tag{7.42}$$

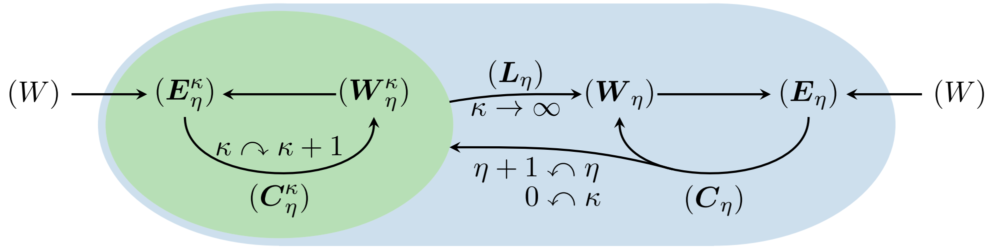

Figure 7.3: Roadmap for the analysis of the abstract fully discrete scheme with the fully implicit midpoint rule.

and $\boldsymbol{y}_{\eta+1}^0 = 2\underline{\boldsymbol{y}}_{\eta+1/2} - \boldsymbol{y}_\eta$, where we use the notation

$$\boldsymbol{\mathcal{A}}_{\eta+1/2}^k := \boldsymbol{\mathcal{A}}(\boldsymbol{y}_{\eta+1/2}^k), \qquad \boldsymbol{\mathcal{F}}_{\eta+1/2}^k := \boldsymbol{\mathcal{F}}(t_n, \boldsymbol{y}_{\eta+1/2}^k). \tag{7.43}$$

The approximation of the midpoint $\underline{\boldsymbol{y}}_{\eta+1/2}$ is the same as in the definition of the linearly implicit midpoint rule (7.6). Similar to (7.4), we employ the notation

$$\boldsymbol{y}_{\eta+1/2}^k = \frac{\boldsymbol{y}_{\eta+1}^k + \boldsymbol{y}_\eta}{2}, \qquad k \geq 0. \tag{7.44}$$

If the sequence given by (7.42) has a fixed point, this is the next approximation $\boldsymbol{y}_{\eta+1}$ of the fully implicit midpoint rule.

Note that in this section we use the same radii as introduced in Section 7.1.1, cf. Figure 7.1.

Roadmap to prove wellposedness and error bounds

The idea for the analysis of the full discretization of quasilinear wave-type problems with the fully implicit midpoint rule is essentially the same as for the linearly implicit midpoint rule. However, the proof of the wellposedness of the scheme is more involved, as we consider a fixed-point iteration to prove existence of the next approximation. Nevertheless, we again assume that the first $\eta \geq 0$ steps are well defined. Then, the roadmap consists of the following steps.

(W) From Assumption 3.2, we have wellposedness of the continuous quasilinear Cauchy problem (3.3). In particular, we get that there is a unique solution y of (3.3), which satisfies $\|y\|_{\mathcal{Y}} < R$ uniformly on J_T. The radius R is given by Assumption 3.1.

$(\boldsymbol{W}_\eta^\kappa)$ Based on the assumption that the first $\eta \geq 0$ steps of the linearly implicit midpoint rule and the first $\kappa \geq 0$ steps of the fixed-point iteration are well defined, we prove in Lemma 7.11 that there exists a range of time steps $\tau \in (0, \tau_{0,\mathrm{w}})$ such that the next approximation $\boldsymbol{y}_{\eta+1}^{\kappa+1}$ of the fixed-point iteration exists and satisfies $\|\boldsymbol{y}_{\eta+1}^{\kappa+1}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}$, using the corresponding result for the linearly implicit midpoint rule.

$(\boldsymbol{E}_\eta^\kappa)$ In Lemma 7.12 we prove an error estimate of the fixed-point iteration, i.e., we prove a bound for the error $\boldsymbol{e}_{\eta+1}^{\kappa+1} = y(t_{\eta+1}) - \boldsymbol{y}_{\eta+1}^{\kappa+1}$ based on the error of the previous approximations $\boldsymbol{e}_{\eta+1}^\kappa$ and $\boldsymbol{e}_\eta$.

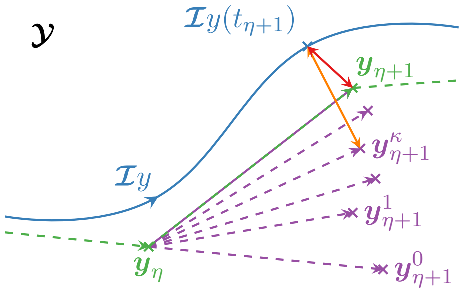

Figure 7.4: Illustration of the fixed-point iteration.

($\boldsymbol{C}_\eta^\kappa$) If for $k \leq \kappa$ the iterates satisfy $\|\boldsymbol{y}_{\eta+1}^k - \boldsymbol{\mathcal{I}}y(t_{\eta+1})\|_{\boldsymbol{\mathcal{Y}}} \to 0$ uniformly in k for $\tau, h \to 0$ under the step size restriction (7.14), we also obtain $\|\boldsymbol{y}_{\eta+1}^{\kappa+1} - \boldsymbol{\mathcal{I}}y(t_{\eta+1})\|_{\boldsymbol{\mathcal{Y}}} \to 0$ using the inverse estimate (4.1). This shows that step $\kappa+1$ of the fixed-point iteration is well defined. Thus, we proceed with ($\boldsymbol{W}_\eta^{\kappa+1}$).

($\boldsymbol{L}_\eta$) In Lemma 7.13, we prove by induction that the fixed-point iteration is well defined, as we employ alternately ($\boldsymbol{W}_\eta^\kappa$), ($\boldsymbol{E}_\eta^\kappa$), and ($\boldsymbol{C}_\eta^\kappa$). Further, we show in Lemma 7.14 that (7.42) defines a Cauchy sequence, which is convergent in the weaker space $\boldsymbol{\mathcal{X}}$.

($\boldsymbol{W}_\eta$) We further show in Lemma 7.14 that the limit $\boldsymbol{y}_{\eta+1}$ of the Cauchy sequence satisfies $\|\boldsymbol{y}_{\eta+1}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}$. Hence, this is the next iterate of the fully implicit midpoint rule.

($\boldsymbol{E}_\eta$) In Lemma 7.15 we show that the fully implicit midpoint rule is stable, i.e., we bound the error $\boldsymbol{e}_{\eta+1} = y(t_{\eta+1}) - \boldsymbol{y}_{\eta+1}$ based on the errors of the previous iterate $\boldsymbol{e}_\eta$.

($\boldsymbol{C}_\eta$) If for $n \leq \eta$ the approximations $\boldsymbol{y}_n$ satisfy $\|\boldsymbol{y}_n - \boldsymbol{\mathcal{I}}y(t_n)\|_{\boldsymbol{\mathcal{Y}}} \to 0$ uniformly in n for $\tau, h \to 0$ under the step size restriction (7.14), we also obtain $\|\boldsymbol{y}_{\eta+1} - \boldsymbol{\mathcal{I}}y(t_{\eta+1})\|_{\boldsymbol{\mathcal{Y}}} \to 0$ using the inverse estimate (4.1). This shows that step $\eta+1$ of the fully implicit midpoint rule is well defined. Thus, we proceed with ($\boldsymbol{W}_{\eta+1}^0$).

Overall, we show Theorem 7.3 by a nested induction, as we alternately prove ($\boldsymbol{W}_\eta$), ($\boldsymbol{E}_\eta$), and ($\boldsymbol{C}_\eta$) to analyze the fully implicit midpoint rule. Further, we also show ($\boldsymbol{W}_\eta$) by induction, i.e., we alternately prove ($\boldsymbol{W}_\eta^\kappa$), ($\boldsymbol{E}_\eta^\kappa$), and ($\boldsymbol{C}_\eta^\kappa$) to show wellposedness of the fixed-point iteration, and finally ($\boldsymbol{L}_\eta$) to prove convergence. This approach is also illustrated in Figure 7.3, where the analysis of the fully implicit midpoint rule is indicated by the blue ellipse. The analysis of the fixed-point iteration is characterized by the green ellipse.

Furthermore, Figure 7.4 is an illustration of the fixed-point iteration, where the interpolation of the solution to the continuous problem is colored in blue and the last iterate $\boldsymbol{y}_\eta$ of the fully implicit midpoint rule is indicated in green. Observe that the iterates of the fixed-point iteration $\boldsymbol{y}_{\eta+1}^k$, which approximate the next iterate $\boldsymbol{y}_{\eta+1}$ of the fully implicit midpoint rule, are plotted in purple. The error $\boldsymbol{e}_{\eta+1}^\kappa$ of the iterate $\boldsymbol{y}_{\eta+1}^\kappa$ of the fixed-point iteration is indicated in orange, whereas the error $\boldsymbol{e}_{\eta+1}$ of the next iterate $\boldsymbol{y}_{\eta+1}$ of the fully implicit midpoint rule is colored in red. Finally, the dashed green line on the right indicates the next but one step of the fully implicit scheme.

Remark 7.10. *Although the introduction of the fixed-point iteration is quite similar to the proof of Banach's fixed-point theorem, the theorem is not directly applicable in our case, as it does not allow for the differentiated treatment of the weaker space* $\boldsymbol{\mathcal{X}}$ *and the stronger space* $\boldsymbol{\mathcal{Y}}$*. Hence, we would need to prove contractivity of* $\boldsymbol{y}_{\eta+1}^{\kappa} \mapsto \boldsymbol{y}_{\eta+1}^{\kappa+1}$ *as a self-mapping on* $B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$*. As we have seen in the proof of Lemma 7.7, this can only be done provided a severe restriction on the time step. However, for the fully implicit midpoint rule, this restriction can not be relaxed afterwards, as this is a nonlinear scheme.*

In the first lemma, we start by proving the wellposedness of one step of the fixed-point iteration.

Lemma 7.11. *For* $0 \le \eta < N$ *and* $\kappa \ge 0$ *fixed assume*

$$\|\boldsymbol{y}_\eta\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}} R), \qquad \|\boldsymbol{\mathcal{A}}_\eta \boldsymbol{y}_\eta\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}} R^{\mathcal{A}}), \tag{7.45}$$

and

$$\|\boldsymbol{y}_{\eta+1/2}^{\kappa}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}. \tag{7.46}$$

There exists $\tau_{0,\mathrm{w}}^{\eta,\kappa} > 0$*, which may depend on the space discretization parameter* h*, such that the fixed-point iteration (7.42) has for all* $\tau < \tau_{0,\mathrm{w}}^{\eta,\kappa}$ *a unique solution* $\boldsymbol{y}_{\eta+1}^{\kappa+1} \in \boldsymbol{\mathcal{X}}$*, which satisfies*

$$\|\boldsymbol{y}_{\eta+1/2}^{\kappa+1}\|_{\boldsymbol{\mathcal{Y}}}, \|\boldsymbol{y}_{\eta+1}^{\kappa+1}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}, \qquad \|\boldsymbol{\mathcal{A}}_{\eta+1/2}^{\kappa} \boldsymbol{y}_{\eta+1/2}^{\kappa+1}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}^{\boldsymbol{\mathcal{A}}}. \tag{7.47}$$

Proof. The result follows from Lemma 7.7 for $\boldsymbol{y}_{\eta+1/2}^{\kappa}$ and $\boldsymbol{y}_{\eta+1}^{\kappa+1}$ instead of $\underline{\boldsymbol{y}}_{\eta+1/2}$ and $\boldsymbol{y}_{\eta+1}$, respectively. □

Based on the notation $\boldsymbol{e}_{\eta+1}^{k} = \boldsymbol{\mathcal{J}} \widetilde{y}_{\eta+1} - \boldsymbol{y}_{\eta+1}^{k}$, we derive equivalently to (7.10) from (7.42) the error equation for the fixed-point iteration

$$\boldsymbol{e}_{\eta+1}^{k+1} = \boldsymbol{e}_\eta + \tau \widetilde{\boldsymbol{\mathcal{A}}}_{\eta+1/2} \boldsymbol{e}_{\eta+1/2}^{k} + \tau \boldsymbol{g}_{\eta+1}^{\mathrm{FI},\,k+1}, \qquad k \ge 0,$$

with right-hand side

$$\boldsymbol{g}_{\eta+1}^{\mathrm{FI},\,k+1} = (\widetilde{\boldsymbol{\mathcal{A}}}_{\eta+1/2} - \boldsymbol{\mathcal{A}}_{\eta+1/2}^{k}) \boldsymbol{y}_{\eta+1/2}^{k+1} - \boldsymbol{\mathcal{F}}_{\eta+1/2}^{k} + \tfrac{1}{\tau} \boldsymbol{\mathcal{J}}(\widetilde{y}_{\eta+1} - \widetilde{y}_\eta) - \widetilde{\boldsymbol{\mathcal{A}}}_{\eta+1/2} \boldsymbol{\mathcal{J}}(\widehat{\delta}_{n+1/2} - \widetilde{y}_{n+1/2}).$$

Using this representation of the error, in the next lemma we show an error recursion for one step of the fixed-point iteration.

Lemma 7.12. *Let Assumption 7.2 be true. If the assumptions of Lemma 7.11 are satisfied for* $0 \le \eta < N$ *and* $\kappa \ge 0$ *fixed, we get for* $\tau < \tau_{0,\mathrm{w}}^{\eta,\kappa}$ *the bound*

$$\begin{aligned} \|\boldsymbol{e}_{\eta+1}^{\kappa+1}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1})}^2 &\le (1 + C\tau) \|\boldsymbol{e}_\eta\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_\eta)}^2 + C_{\underline{e}}' \tau \|\boldsymbol{e}_{\eta+1}^{\kappa}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1})}^2 \\ &\quad + C\tau^5 \Big(\sup_{[t_\eta, t_{\eta+1}]} \|\partial_t^3 y\|_{\mathcal{X}}^2 + \sup_{[t_{\eta-1}, t_{\eta+1}]} \|\partial_t^2 y\|_{\mathcal{Y}}^2 \Big) + C\tau \|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}}) \widetilde{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{X}}}^2 \\ &\quad + C\tau \Big(\sup_{[t_\eta, t_{\eta+1}]} \|\mathcal{R}_\Lambda(y) \partial_t y\|_{\boldsymbol{\mathcal{X}}}^2 + \sup_{[t_\eta, t_{\eta+1}]} \|\mathcal{R}_{\mathrm{A}} y\|_{\boldsymbol{\mathcal{X}}}^2 + \sup_{[t_\eta, t_{\eta+1}]} \|\mathcal{R}_{\mathrm{F}}(\cdot, y)\|_{\boldsymbol{\mathcal{X}}}^2 \Big), \end{aligned} \tag{7.48}$$

with constants $C, C_{\underline{e}}' > 0$ *independent of* η*,* κ*,* h*, and* τ*.*

Proof. Using Lemma 7.8 with the same replacements as in the previous proof, i.e., we take $\boldsymbol{y}_{\eta+1/2}^{\kappa}$ and $\boldsymbol{y}_{\eta+1}^{\kappa+1}$ instead of $\underline{\boldsymbol{y}}_{\eta+1/2}$ and $\boldsymbol{y}_{\eta+1}$, respectively, directly yields

$$\begin{aligned}\|\boldsymbol{e}_{\eta+1}^{\kappa+1}\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1})}^2 &\leq (1+C\tau)\|\boldsymbol{e}_\eta\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_\eta)}^2 + C_{\underline{e}}\tau\|\boldsymbol{\mathcal{J}}\widetilde{y}_{\eta+1/2} - \boldsymbol{y}_{\eta+1/2}^{\kappa}\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1/2})}^2 \\ &\quad + C\tau^5\Big(\sup_{[t_\eta,t_{\eta+1}]}\|\partial_t^3 y\|_{\mathcal{X}}^2 + \sup_{[t_{\eta-1},t_{\eta+1}]}\|\partial_t^2 y\|_{\mathcal{Y}}^2\Big) + C\tau\|(\boldsymbol{\mathcal{I}}-\boldsymbol{\mathcal{J}})\widetilde{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{X}}}^2 \\ &\quad + C\tau\Big(\sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_\Lambda(y)\partial_t y\|_{\boldsymbol{\mathcal{X}}}^2 + \sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_{\mathrm{A}} y\|_{\boldsymbol{\mathcal{X}}}^2 + \sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_{\mathrm{F}}(\cdot,y)\|_{\boldsymbol{\mathcal{X}}}^2\Big).\end{aligned}$$

It remains to derive a bound for $\|\boldsymbol{y}_{\eta+1/2}^{\kappa} - \boldsymbol{\mathcal{J}}\widetilde{y}_{\eta+1/2}\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1/2})}^2$. With similar arguments as in (7.39), we derive

$$\begin{aligned}\|\boldsymbol{\mathcal{J}}\widetilde{y}_{\eta+1/2} - \boldsymbol{y}_{\eta+1/2}^{\kappa}\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1/2})} &\leq \tfrac{1}{2}\|\boldsymbol{\mathcal{J}}\widetilde{y}_{\eta+1} - \boldsymbol{y}_{\eta+1}^{\kappa}\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1/2})} + \tfrac{1}{2}\|\boldsymbol{\mathcal{J}}\widetilde{y}_\eta - \boldsymbol{y}_\eta\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1/2})} \\ &\quad + \tfrac{1}{2}\|\boldsymbol{\mathcal{J}}(2\widetilde{y}_{\eta+1/2} - \widetilde{y}_{\eta+1} - \widetilde{y}_\eta)\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1/2})} \\ &\leq \tfrac{1}{2}\|\boldsymbol{e}_{\eta+1}^{\kappa}\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1/2})} + \tfrac{1}{2}\|\boldsymbol{e}_\eta\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1/2})} + C\tau^2\sup_{[t_\eta,t_{\eta+1}]}\|\partial_t^2 y\|_{\mathcal{Y}},\end{aligned}$$

where we used again Taylor's theorem in the last step. Using the norm equivalence (4.11) as in (7.35) finally yields the error estimate (7.48), which concludes the proof. □

With these two lemmas at hand, we are now able to show wellposedness of the fixed-point iteration.

Lemma 7.13. *Let Assumptions 4.22 and 7.2 be true. Furthermore, for $0 \leq \eta < N$ fixed we assume (7.45) and that errors of the previous iterates computed with the fully implicit midpoint rule satisfy*

$$C_{\max}(h)\|\boldsymbol{e}_{\eta-1}\|_{\boldsymbol{\mathcal{X}}}, C_{\max}(h)\|\boldsymbol{e}_\eta\|_{\boldsymbol{\mathcal{X}}} \to 0 \tag{7.49}$$

uniformly in η for $h, \tau \to 0$ under the step size restriction (7.14). Then, there exist $\tau_0, h_0 > 0$ such that the fixed-point iteration (7.42) is wellposed for all $h < h_0$ and $\tau < \tau_0$ under the step size restriction (7.14), i.e., the iterates satisfy

$$\|\boldsymbol{y}_{\eta+1}^k\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R), \qquad \|\boldsymbol{\mathcal{A}}_{\eta+1}^k\boldsymbol{y}_{\eta+1}^k\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R^{\mathcal{A}}), \qquad k \geq 0. \tag{7.50}$$

Proof. Let $0 \leq \eta < N$ be fixed. The proof is done by induction over $k \geq 0$, where we alternately prove the error estimate

$$\|\boldsymbol{e}_{\eta+1}^k\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1})}^2 \leq (C_{\underline{e}}'\tau)^k\|\boldsymbol{e}_{\eta+1}^0\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1})}^2 + C_{\mathrm{fp}}(\tau_0,h_0)^2\sum_{i=1}^k (C_{\underline{e}}'\tau)^{i-1}, \qquad k \geq 0, \tag{7.51}$$

and derive the bound (7.50).

Note that for constants $\tau_0 \leq \tau_{0,\mathrm{w}}^{\eta,\kappa}$ and $h_0 > 0$, which are fixed later in the proof, the term $C_{\mathrm{fp}}(\tau_0,h_0)$ appearing in (7.51) is chosen as the minimal constant, which satisfies

$$\begin{aligned}C_{\mathrm{fp}}(\tau_0,h_0)^2 &\geq (1+C\tau)\|\boldsymbol{e}_\eta\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_\eta)}^2 + C\tau\|(\boldsymbol{\mathcal{I}}-\boldsymbol{\mathcal{J}})\widetilde{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{X}}}^2 \\ &\quad + C\tau^5\Big(\sup_{[t_\eta,t_{\eta+1}]}\|\partial_t^3 y\|_{\mathcal{X}}^2 + \sup_{[t_{\eta-1},t_{\eta+1}]}\|\partial_t^2 y\|_{\mathcal{Y}}^2\Big) + C\tau\Big(\sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_\Lambda(y)y\|_{\boldsymbol{\mathcal{X}}}^2 \\ &\quad + \sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_\Lambda(y)\partial_t y\|_{\boldsymbol{\mathcal{X}}}^2 + \sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_{\mathrm{A}} y\|_{\boldsymbol{\mathcal{X}}}^2 + \sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_{\mathrm{F}}(\cdot,y)\|_{\boldsymbol{\mathcal{X}}}^2\Big),\end{aligned}$$

for all $\tau < \tau_0$ and $h < h_0$ under the step size restriction (7.14). For τ_0 and h_0 sufficiently small, this yields in fact a finite constant, as all terms on the right-hand side tend to 0 for $h, \tau \to 0$ under the step size restriction (7.14). In addition, Assumption 4.22 even yields

$$C_{\max}(h) C_{\mathrm{fp}}(\tau_0, h_0) \to 0, \tag{7.52}$$

for $\tau_0, h_0 \to 0$. Using the estimate (7.51), we then derive the bound (7.50).

For the induction base $k = 0$, there is nothing to show, as (7.51) is trivially satisfied. Also, the first step of the fixed-point iteration (7.42) corresponds to one step of the linearly implicit midpoint rule with initial value $\boldsymbol{y}_\eta$. Thus, (7.50) follows directly from the analysis for the linearly implicit midpoint rule.

For the induction hypothesis, we assume that (7.50) and (7.51) hold up to some $k = \kappa \geq 0$ arbitrary but fixed. To prove the induction step, we first use Lemma 7.11 to get the existence of the next iterates. This is applicable, since (7.46) follows directly from (7.45) and the induction hypothesis. Next, we employ the error estimate (7.48) to show the error bound

$$\|\boldsymbol{e}_{\eta+1}^{\kappa+1}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1})} \leq C'_{\underline{\boldsymbol{e}}} \tau \|\boldsymbol{e}_{\eta+1}^{\kappa}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1})} + C_{\mathrm{fp}}(\tau_0, h_0)^2.$$

Using the induction hypothesis to replace $\|\boldsymbol{e}_{\eta+1}^{\kappa}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1})}$, this proves (7.51) for $k = \kappa+1$. Based on this estimate, we now prove (7.50).

Since $\boldsymbol{e}_{\eta+1}^1$ corresponds to the error after one further step with the linearly implicit midpoint rule with initial value $\boldsymbol{y}_\eta$, the error estimate (7.28) together with the bound (7.39) yields

$$\|\boldsymbol{e}_{\eta+1}^{0}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1})} \leq C_{\mathrm{fp}}(\tau_0, h_0)^2 + C\tau \big(\|\boldsymbol{e}_\eta\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} + \|\boldsymbol{e}_{\eta-1}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} + \tau^4 \sup_{[t_{\eta-1}, t_{\eta+1/2}]} \|\partial_t^2 y\|^2_{\mathcal{Y}}\big).$$

Furthermore, since the constant $C'_{\underline{\boldsymbol{e}}}$ is in particular independent of h, τ, and κ, we get for $\tau < \frac{1}{C'_{\underline{\boldsymbol{e}}}}$ that the sum appearing in (7.51) is a geometric sum. Using the inverse estimate (4.1) together with the norm equivalence (4.10), these bounds yield

$$\|\boldsymbol{e}_{\eta+1}^{\kappa+1}\|^2_{\boldsymbol{\mathcal{Y}}} \leq C \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h) \tau \big(\|\boldsymbol{e}_\eta\|_{\boldsymbol{\mathcal{X}}} + \|\boldsymbol{e}_{\eta-1}\|_{\boldsymbol{\mathcal{X}}} + \tau^4 \sup_{[t_{\eta-1}, t_{\eta+1}]} \|\partial_t^2 y\|_{\mathcal{Y}}\big) + \frac{C C_{\mathrm{fp}}(\tau_0, h_0)^2}{1 - C'_{\underline{\boldsymbol{e}}} \tau}.$$

Using (7.49) and (7.52), we get that all terms on the right side of the inequality tend to 0 uniformly in η for $\tau_0, h_0 \to 0$ with $\tau < \tau_0$ and $h < h_0$ satisfying the step size restriction (7.14). In particular, we take $h_0, \tau_0 > 0$ with $\tau_0 < \frac{1}{C'_{\underline{\boldsymbol{e}}}}$ such that the assumptions of Lemma 4.24 are satisfied, which yields (7.50). □

Up to now, we have shown that the sequence given by (7.42) is well defined. In the next lemma, we prove the convergence of these sequences in $\boldsymbol{\mathcal{X}}$. Note that, since we are able to show that the limit is even contained in $B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$, this proves wellposedness of one step of the fully implicit midpoint rule.

Lemma 7.14. *Let $0 \leq \eta < N$ fixed. Under the assumptions of* Lemma 7.13*, the sequence defined by* (7.42) *converges in $\boldsymbol{\mathcal{X}}$ with limit $\boldsymbol{y}_{\eta+1} \in \boldsymbol{\mathcal{X}}$, which satisfies*

$$\|\boldsymbol{y}_{\eta+1}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}, \qquad \|\boldsymbol{\mathcal{A}}_{\eta+1} \boldsymbol{y}_{\eta+1}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}^{\boldsymbol{\mathcal{A}}}.$$

Proof. To prove the statement, we first prove that the sequence is a Cauchy sequence in $\boldsymbol{\mathcal{X}}$. To do so, we bound the difference between two consecutive elements. Next, we use this estimate to bound the difference between arbitrary elements.

1. For $k \geq 1$, (7.42) yields

$$\boldsymbol{y}_{\eta+1}^{k+1}-\boldsymbol{y}_{\eta+1}^{k}-\tfrac{\tau}{2}\boldsymbol{\mathcal{A}}_{\eta+1/2}^{k}(\boldsymbol{y}_{\eta+1}^{k+1}-\boldsymbol{y}_{\eta+1}^{k}) = \tau(\boldsymbol{\mathcal{A}}_{\eta+1/2}^{k}-\boldsymbol{\mathcal{A}}_{\eta+1/2}^{k-1})\boldsymbol{y}_{\eta+1/2}^{k} + \tau(\boldsymbol{\mathcal{F}}_{\eta+1/2}^{k}-\boldsymbol{\mathcal{F}}_{\eta+1/2}^{k-1}).$$

   Taking the $\mathbf{\Lambda}(\boldsymbol{y}_{\eta+1/2}^{k})$ inner product with $\boldsymbol{y}_{\eta+1}^{k+1}-\boldsymbol{y}_{\eta+1}^{k}$, and using the dissipativity of $\mathbf{A}$, we derive the bound

$$\begin{aligned}\|\boldsymbol{y}_{\eta+1}^{k+1}-\boldsymbol{y}_{\eta+1}^{k}\|_{\mathbf{\Lambda}(\boldsymbol{y}_{\eta+1/2}^{k+1})}^2 \leq{}& \tau\left((\boldsymbol{\mathcal{A}}_{\eta+1/2}^{k}-\boldsymbol{\mathcal{A}}_{\eta+1/2}^{k-1})\boldsymbol{y}_{\eta+1/2}^{k} \mid \boldsymbol{y}_{\eta+1}^{k+1}-\boldsymbol{y}_{\eta+1}^{k}\right)_{\mathbf{\Lambda}(\boldsymbol{y}_{\eta+1/2}^{k+1})}\\ &+\tau\left(\boldsymbol{\mathcal{F}}_{\eta+1/2}^{k}-\boldsymbol{\mathcal{F}}_{\eta+1/2}^{k-1} \mid \boldsymbol{y}_{\eta+1}^{k+1}-\boldsymbol{y}_{\eta+1}^{k}\right)_{\mathbf{\Lambda}(\boldsymbol{y}_{\eta+1/2}^{k+1})}.\end{aligned}$$

   Hence, using the Cauchy–Schwarz inequality yields

$$\begin{aligned}\|\boldsymbol{y}_{\eta+1}^{k+1}-\boldsymbol{y}_{\eta+1}^{k}\|_{\mathbf{\Lambda}(\boldsymbol{y}_{\eta+1/2}^{k+1})} \leq{}& \tau\|(\boldsymbol{\mathcal{A}}_{\eta+1/2}^{k}-\boldsymbol{\mathcal{A}}_{\eta+1/2}^{k-1})\boldsymbol{y}_{\eta+1/2}^{k}\|_{\mathbf{\Lambda}(\boldsymbol{y}_{\eta+1/2}^{k+1})}\\ &+\tau\|\boldsymbol{\mathcal{F}}_{\eta+1/2}^{k}-\boldsymbol{\mathcal{F}}_{\eta+1/2}^{k-1}\|_{\mathbf{\Lambda}(\boldsymbol{y}_{\eta+1/2}^{k+1})}.\end{aligned}$$

   Based on the norm equivalence (4.10), the Lipschitz continuity (4.14) and (4.15) of both operators, as well as the bound (7.47) on all iterates, we further deduce

$$\begin{aligned}\|\boldsymbol{y}_{\eta+1}^{k+1}-\boldsymbol{y}_{\eta+1}^{k}\|_{\boldsymbol{\mathcal{X}}} &\leq \tau \boldsymbol{c}_{\mathbf{\Lambda}}^{-\frac{1}{2}}\boldsymbol{C}_{\mathbf{\Lambda}}^{\frac{1}{2}}(\boldsymbol{L}_{\boldsymbol{\mathcal{A}}}\|\boldsymbol{\mathcal{A}}_{\eta+1/2}^{k}\boldsymbol{y}_{\eta+1/2}^{k}\|_{\boldsymbol{\mathcal{Y}}}+\boldsymbol{L}_{\boldsymbol{\mathcal{F}}})\|\boldsymbol{y}_{\eta+1}^{k}-\boldsymbol{y}_{\eta+1}^{k-1}\|_{\boldsymbol{\mathcal{X}}}\\ &\leq \tau \boldsymbol{c}_{\mathbf{\Lambda}}^{-\frac{1}{2}}\boldsymbol{C}_{\mathbf{\Lambda}}^{\frac{1}{2}}(\boldsymbol{L}_{\boldsymbol{\mathcal{A}}}\boldsymbol{R}^{\boldsymbol{\mathcal{A}}}+\boldsymbol{L}_{\boldsymbol{\mathcal{F}}})\|\boldsymbol{y}_{\eta+1}^{k}-\boldsymbol{y}_{\eta+1}^{k-1}\|_{\boldsymbol{\mathcal{X}}}.\end{aligned}$$

   Hence, there exists $\tau_{0,\mathrm{c}} > 0$ such that for all $\tau < \tau_{0,\mathrm{c}}$, we have

$$\|\boldsymbol{y}_{\eta+1}^{k+1}-\boldsymbol{y}_{\eta+1}^{k}\|_{\boldsymbol{\mathcal{X}}} \leq \varepsilon_{\mathrm{c}}\|\boldsymbol{y}_{\eta+1}^{k}-\boldsymbol{y}_{\eta+1}^{k-1}\|_{\boldsymbol{\mathcal{X}}}$$

   with

$$\varepsilon_{\mathrm{c}} = \tau_{0,\mathrm{c}}\boldsymbol{c}_{\mathbf{\Lambda}}^{-\frac{1}{2}}\boldsymbol{C}_{\mathbf{\Lambda}}^{\frac{1}{2}}(\boldsymbol{L}_{\boldsymbol{\mathcal{A}}}\boldsymbol{R}^{\boldsymbol{\mathcal{A}}}+\boldsymbol{L}_{\boldsymbol{\mathcal{F}}}) \in (0,1).$$

   Using this argument iteratively, we finally derive

$$\|\boldsymbol{y}_{\eta+1}^{k+1}-\boldsymbol{y}_{\eta+1}^{k}\|_{\boldsymbol{\mathcal{X}}} \leq \varepsilon_{\mathrm{c}}^{k}\|\boldsymbol{y}_{\eta+1}^{1}-\boldsymbol{y}_{\eta+1}^{0}\|_{\boldsymbol{\mathcal{X}}}. \tag{7.53}$$

2. Let $\ell > m \geq 1$. The triangle inequality together with (7.53) yields

$$\begin{aligned}\|\boldsymbol{y}_{\eta+1}^{\ell}-\boldsymbol{y}_{\eta+1}^{m}\|_{\boldsymbol{\mathcal{X}}} &\leq \|\boldsymbol{y}_{\eta+1}^{\ell}-\boldsymbol{y}_{\eta+1}^{\ell-1}\|_{\boldsymbol{\mathcal{X}}}+\cdots+\|\boldsymbol{y}_{\eta+1}^{m+1}-\boldsymbol{y}_{\eta+1}^{m}\|_{\boldsymbol{\mathcal{X}}}\\ &\leq \varepsilon_{\mathrm{c}}^{m}(\varepsilon_{\mathrm{c}}^{\ell-m-1}+\cdots+\varepsilon_{\mathrm{c}}+1)\|\boldsymbol{y}_{\eta+1}^{1}-\boldsymbol{y}_{\eta+1}^{0}\|_{\boldsymbol{\mathcal{X}}}.\end{aligned}$$

   Finally, since the sum is a geometric sum, we get

$$\|\boldsymbol{y}_{\eta+1}^{\ell}-\boldsymbol{y}_{\eta+1}^{m}\|_{\boldsymbol{\mathcal{X}}} \leq \frac{\varepsilon_{\mathrm{c}}^{m}}{1-\varepsilon_{\mathrm{c}}}\|\boldsymbol{y}_{\eta+1}^{1}-\boldsymbol{y}_{\eta+1}^{0}\|_{\boldsymbol{\mathcal{X}}},$$

   which proves that the sequence of iterates defined by (7.42) is a Cauchy sequence.

Since $\boldsymbol{\mathcal{X}}$ is a complete space, the Cauchy sequence is convergent with limit $\boldsymbol{y}_{\eta+1} \in \boldsymbol{\mathcal{X}}$. Hence, there exists $k_0 \in \mathbb{N}$, which may depend on τ and h such that both

$$\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\|\boldsymbol{y}_{\eta+1} - \boldsymbol{y}_{\eta+1}^k\|_{\boldsymbol{\mathcal{X}}} \le \tfrac{1}{2}(\boldsymbol{R} - \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R) \tag{7.54}$$

and

$$\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\Big(\boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}\boldsymbol{C}_{\mathbf{A}}(h) + \tfrac{1}{2}\boldsymbol{L}_{\boldsymbol{\mathcal{A}}}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R^{\mathcal{A}})\Big)\|\boldsymbol{y}_{\eta+1} - \boldsymbol{y}_{\eta+1}^k\|_{\boldsymbol{\mathcal{X}}} \le \tfrac{1}{2}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} - \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R^{\mathcal{A}}) \tag{7.55}$$

hold for all $k \ge k_0$. On the one hand, the inverse estimate (4.1), (7.54), and (7.50) yield the bound

$$\|\boldsymbol{y}_{\eta+1}\|_{\boldsymbol{\mathcal{Y}}} \le \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\|\boldsymbol{y}_{\eta+1} - \boldsymbol{y}_{\eta+1}^k\|_{\boldsymbol{\mathcal{X}}} + \|\boldsymbol{y}_{\eta+1}^k\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}.$$

On the other hand, we first compute with the boundedness (4.16) and (4.29) of $\boldsymbol{\Lambda}^{-1}$ and $\mathbf{A}$, respectively,

$$\|\boldsymbol{\mathcal{A}}_{\eta+1}(\boldsymbol{y}_{\eta+1} - \boldsymbol{y}_{\eta+1}^k)\|_{\boldsymbol{\mathcal{X}}} \le \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}\boldsymbol{C}_{\mathbf{A}}(h)\|\boldsymbol{y}_{\eta+1} - \boldsymbol{y}_{\eta+1}^k\|_{\boldsymbol{\mathcal{X}}}.$$

We further use the Lipschitz continuity of $\boldsymbol{\mathcal{A}}$ (4.14) together with (7.50) to show

$$\begin{aligned}\|(\boldsymbol{\mathcal{A}}_{\eta+1} - \boldsymbol{\mathcal{A}}_{\eta+1}^k)\boldsymbol{y}_{\eta+1}^k\|_{\boldsymbol{\mathcal{X}}} &\le \boldsymbol{L}_{\boldsymbol{\mathcal{A}}}\|\boldsymbol{\mathcal{A}}_{\eta+1}^k\boldsymbol{y}_{\eta+1}^k\|_{\boldsymbol{\mathcal{Y}}}\|\boldsymbol{y}_{\eta+1} - \boldsymbol{y}_{\eta+1}^k\|_{\boldsymbol{\mathcal{X}}}\\ &\le \tfrac{1}{2}\boldsymbol{L}_{\boldsymbol{\mathcal{A}}}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}R^{\mathcal{A}})\|\boldsymbol{y}_{\eta+1} - \boldsymbol{y}_{\eta+1}^k\|_{\boldsymbol{\mathcal{X}}}.\end{aligned}$$

Hence, we finally use these bounds together with the inverse estimate (4.1), (7.50), and (7.55) to derive

$$\|\boldsymbol{\mathcal{A}}_{\eta+1}\boldsymbol{y}_{\eta+1}\|_{\boldsymbol{\mathcal{Y}}} \le \|\boldsymbol{\mathcal{A}}_{\eta+1}(\boldsymbol{y}_{\eta+1} - \boldsymbol{y}_{\eta+1}^k)\|_{\boldsymbol{\mathcal{Y}}} + \|(\boldsymbol{\mathcal{A}}_{\eta+1} - \boldsymbol{\mathcal{A}}_{\eta+1}^k)\boldsymbol{y}_{\eta+1}^k\|_{\boldsymbol{\mathcal{Y}}} + \|\boldsymbol{\mathcal{A}}_{\eta+1}\boldsymbol{y}_{\eta+1}^k\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}^{\boldsymbol{\mathcal{A}}},$$

which proves the statement. □

We further present an error bound based on Lemma 7.8 for the fully implicit midpoint rule.

Lemma 7.15. *Let Assumption 7.2 be true. If the assumptions of Lemma 7.7 are satisfied for $0 \le \eta < N$ fixed, there exists $\tau_{0,\mathrm{e}} > 0$ such that the error of the fully implicit midpoint rule satisfies for $\tau < \min\{\tau_{0,\mathrm{w}}^{\eta,\kappa}, \tau_{0,\mathrm{e}}\}$ the bound*

$$\begin{aligned}\|\boldsymbol{e}_{\eta+1}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1})} &\le (1 + C\tau)\|\boldsymbol{e}_\eta\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_\eta)} + C\tau^5\Big(\sup_{[t_\eta,t_{\eta+1}]}\|\partial_t^3 y\|^2_{\mathcal{X}} + \sup_{[t_{\eta-1},t_{\eta+1}]}\|\partial_t^2 y\|^2_{\mathcal{Y}}\Big)\\ &\quad + C\tau\Big(\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\widetilde{y}_{\eta+1/2}\|^2_{\boldsymbol{\mathcal{X}}} + \sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_\Lambda(y)\partial_t y\|^2_{\boldsymbol{\mathcal{X}}} + \sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_{\mathrm{A}}y\|^2_{\boldsymbol{\mathcal{X}}} + \sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_{\mathrm{F}}(\cdot,y)\|^2_{\boldsymbol{\mathcal{X}}}\Big).\end{aligned}$$

Proof. To prove the result, we use a similar argumentation as in the proof of Lemma 7.8, but for the right-hand side of the fully implicit scheme defined in (7.11). To be more precise, we first obtain similarly to (7.34) for the right-hand side of the fully implicit scheme

$$\begin{aligned}\|\boldsymbol{g}_{\eta+1}^{\mathrm{FI}}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} \le C\big(&\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\widetilde{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{X}}} + \|\boldsymbol{\mathcal{J}}\widetilde{y}_{\eta+1/2} - \boldsymbol{y}_{\eta+1/2}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})}\\ &+ \sup_{[t_\eta,t_{\eta+1}]}\|\mathcal{R}_\Lambda(\widetilde{y}_{\eta+1/2})\partial_t y\|_{\boldsymbol{\mathcal{X}}} + \|\mathcal{R}_{\mathrm{A}}(\widetilde{y}_{\eta+1} + \widetilde{y}_\eta)\|_{\boldsymbol{\mathcal{X}}} + \|\mathcal{R}_{\mathrm{F}}(t_{\eta+1/2}, \widetilde{y}_{\eta+1/2})\|_{\boldsymbol{\mathcal{X}}}\\ &+ \|\mathrm{A}\widehat{\delta}_{\eta+1/2}\|_{\mathcal{X}} + \|\tfrac{1}{\tau}\delta_{\eta+1}\|_{\mathcal{X}}\big).\end{aligned}$$

Furthermore, we derive with (7.8), the norm equivalence (4.10), the boundedness (4.17) of $\boldsymbol{\mathcal{J}}$, and the estimate (7.31) the bound

$$\begin{aligned}\|\boldsymbol{\mathcal{J}}\widetilde{y}_{\eta+1/2} - \boldsymbol{y}_{\eta+1/2}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} &\le \|\boldsymbol{\mathcal{J}}\widehat{\delta}_{\eta+1/2}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} + \|\boldsymbol{e}_{\eta+1/2}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} \\ &\le \boldsymbol{C}_{\boldsymbol{\Lambda}}^{\frac{1}{2}}\boldsymbol{C}_{\boldsymbol{\mathcal{J}}}\|\widehat{\delta}_{\eta+1/2}\|_{\boldsymbol{\mathcal{Y}}} + \|\boldsymbol{e}_\eta\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} + \tfrac{\tau}{2}\|\boldsymbol{g}_{\eta+1}^{\mathrm{FI}}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})}.\end{aligned}$$

As $\mathrm{A} \in \mathcal{L}(\mathcal{Y}, \mathcal{X})$, we thus have shown

$$\begin{aligned}\|\boldsymbol{g}_{\eta+1}^{\mathrm{FI}}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} \le C\big(&\|\boldsymbol{e}_\eta\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} + \tau\|\boldsymbol{g}_{\eta+1}^{\mathrm{FI}}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} + \|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\widetilde{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{X}}} \\ &+ \sup_{[t_\eta, t_{\eta+1}]} \|\mathcal{R}_{\Lambda}(\widetilde{y}_{\eta+1/2})\partial_t y\|_{\boldsymbol{\mathcal{X}}} + \|\mathcal{R}_{\mathrm{A}}\widetilde{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{X}}} + \|\mathcal{R}_{\mathrm{F}}(t_{\eta+1/2}, \widetilde{y}_{\eta+1/2})\|_{\boldsymbol{\mathcal{X}}} \\ &+ \|\widehat{\delta}_{n+1/2}\|_{\mathcal{Y}} + \|\tfrac{1}{\tau}\delta_{\eta+1}\|_{\boldsymbol{\mathcal{X}}}\big),\end{aligned}$$

with a constant $C > 0$ independent of τ and h. Hence, there exists $\tau_{0,\mathrm{e}} > 0$ such that $1 - C\tau$ is invertible for all $\tau < \tau_{0,\mathrm{e}}$ and further

$$\begin{aligned}\|\boldsymbol{g}_{\eta+1}^{\mathrm{FI}}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} \le (1 - C\tau)^{-1}C\big(&\|\boldsymbol{e}_\eta\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})} + \|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\widetilde{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{X}}} + \sup_{[t_\eta, t_{\eta+1}]} \|\mathcal{R}_{\Lambda}(\widetilde{y}_{\eta+1/2})\partial_t y\|_{\boldsymbol{\mathcal{X}}} \\ &+ \|\mathcal{R}_{\mathrm{A}}\widetilde{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{X}}} + \|\mathcal{R}_{\mathrm{F}}(t_{\eta+1/2}, \widetilde{y}_{\eta+1/2})\|_{\boldsymbol{\mathcal{X}}} + \|\widehat{\delta}_{n+1/2}\|_{\mathcal{Y}} + \|\tfrac{1}{\tau}\delta_{\eta+1}\|_{\boldsymbol{\mathcal{X}}}\big)\end{aligned}$$

holds for all $\tau < \tau_{0,\mathrm{e}}$. Using this bound in (7.32) together with Lemma 7.6 yields the result. □

Based on the previous lemmas, we are now able to prove the main result of the fully implicit midpoint rule (7.3).

Proof of Theorem 7.3. As for the linearly implicit midpoint rule, this proof is done by induction, i.e., by alternately using Lemma 7.14 to prove existence of the next iterates and Lemma 7.15 to prove the error bound (7.38). Finally, based on Assumption 4.22, the error estimate from Lemma 7.15 together with Lemma 4.24 yields that that the assumptions of Lemma 7.14 are then also satisfied for the next step under the corresponding restrictions on the discretization parameters. □

Finally, we conclude this section with the following remark on how the results obtained in this section can be extended to higher-order time-integration schemes.

Remark 7.16. *As stated in Section 6.1, the implicit midpoint rule is an algebraically stable, coercive Runge–Kutta scheme. Hence, a natural generalization would be to consider also the full discretization with general algebraically stable, coercive Runge–Kutta schemes. For the corresponding linearized schemes of higher order, however, the extrapolation can not only be based on the last two iterates but either on multiple of the previous iterates or also on the inner stages of previous time steps. In both cases, the construction of suitable approximations for the first steps is more involved.*

7.2 Leapfrog scheme

For the full discretization of linear wave-type problems, explicit time-integration schemes like the leapfrog scheme in general have the drawback of being only stable under a step size restriction, whereas implicit schemes are in many applications unconditionally stable. However, as we have seen in the previous sections, implicit schemes like the linearly and fully implicit midpoint rule are for quasilinear wave-type problems also dependent on step size restrictions, which originate from the problem itself. Hence, it is only natural to consider also explicit schemes, which ideally do not deteriorate the step size restriction inherited from the problem, but are computationally cheaper than the implicit schemes considered before. Thus, we present in the following the wellposedness analysis as well as an error estimate for the full discretization based on the leapfrog scheme.

Analogous to the assumptions stated in Section 6.2 for the continuous problem, we assume that they are also inherited by the spatially discrete problem, i.e., we require that for discrete spaces $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{X}}_{\mathcal{V}} \times \boldsymbol{\mathcal{X}}_{\mathcal{H}}$ and $\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{Y}}_{\mathcal{V}} \times \boldsymbol{\mathcal{Y}}_{\mathcal{H}}$, we can rewrite the spatially discrete quasilinear evolution equation (4.12) as

$$\begin{cases} \partial_t \boldsymbol{y}(t) = \boldsymbol{\mathcal{A}}(\boldsymbol{y}(t))\boldsymbol{y}(t) + \boldsymbol{\mathcal{F}}(t, \boldsymbol{y}(t)), \quad \boldsymbol{y} = \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix}, t \in J_T, \\ \boldsymbol{y}(0) = \boldsymbol{y}_0, \end{cases} \tag{7.56}$$

with

$$\boldsymbol{\Lambda}(\boldsymbol{y}) = \begin{pmatrix} \boldsymbol{\Lambda}_{\boldsymbol{\mathcal{V}}}(\boldsymbol{u},\boldsymbol{v}) & 0 \\ 0 & \boldsymbol{\Lambda}_{\boldsymbol{\mathcal{H}}}(\boldsymbol{u},\boldsymbol{v}) \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} 0 & \mathbf{A}_{\boldsymbol{\mathcal{V}}} \\ \mathbf{A}_{\boldsymbol{\mathcal{H}}} & 0 \end{pmatrix}, \quad \mathbf{F}(t,\boldsymbol{y}) = \begin{pmatrix} \mathbf{F}_{\boldsymbol{\mathcal{V}}}(t,\boldsymbol{u},\boldsymbol{v}) \\ \mathbf{F}_{\boldsymbol{\mathcal{H}}}(t,\boldsymbol{u},\boldsymbol{v}) \end{pmatrix}.$$

We further define as in (4.13)

$$\boldsymbol{\mathcal{A}}(\boldsymbol{y}) = \begin{pmatrix} 0 & \boldsymbol{\mathcal{A}}_{\boldsymbol{\mathcal{V}}}(\boldsymbol{u},\boldsymbol{v}) \\ \boldsymbol{\mathcal{A}}_{\boldsymbol{\mathcal{H}}}(\boldsymbol{u},\boldsymbol{v}) & 0 \end{pmatrix} := \begin{pmatrix} 0 & \boldsymbol{\Lambda}_{\boldsymbol{\mathcal{V}}}(\boldsymbol{u},\boldsymbol{v})^{-1}\mathbf{A}_{\boldsymbol{\mathcal{V}}} \\ \boldsymbol{\Lambda}_{\boldsymbol{\mathcal{H}}}(\boldsymbol{u},\boldsymbol{v})^{-1}\mathbf{A}_{\boldsymbol{\mathcal{H}}} & 0 \end{pmatrix}, \ \boldsymbol{y} = \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$$

and

$$\boldsymbol{\mathcal{F}}(t,\boldsymbol{y}) = \begin{pmatrix} \boldsymbol{\mathcal{F}}_{\boldsymbol{\mathcal{V}}}(t,\boldsymbol{u},\boldsymbol{v}) \\ \boldsymbol{\mathcal{F}}_{\boldsymbol{\mathcal{H}}}(t,\boldsymbol{u},\boldsymbol{v}) \end{pmatrix} := \begin{pmatrix} \boldsymbol{\Lambda}_{\boldsymbol{\mathcal{V}}}(\boldsymbol{u},\boldsymbol{v})^{-1}\mathbf{F}_{\boldsymbol{\mathcal{V}}}(t,\boldsymbol{u},\boldsymbol{v}) \\ \boldsymbol{\Lambda}_{\boldsymbol{\mathcal{H}}}(\boldsymbol{u},\boldsymbol{v})^{-1}\mathbf{F}_{\boldsymbol{\mathcal{H}}}(t,\boldsymbol{u},\boldsymbol{v}) \end{pmatrix}, \quad t \in J_T, \ \boldsymbol{y} = \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}).$$

Corresponding to the previous assumptions, we also assume that the operators defined at the end of Section 4.1 have additional structure, i.e., we set

$$\boldsymbol{\mathcal{J}} = \begin{pmatrix} \boldsymbol{\mathcal{J}}_{\mathcal{V}} & 0 \\ 0 & \boldsymbol{\mathcal{J}}_{\mathcal{H}} \end{pmatrix}, \qquad \boldsymbol{\mathcal{L}} = \begin{pmatrix} \boldsymbol{\mathcal{L}}_{\mathcal{V}} & 0 \\ 0 & \boldsymbol{\mathcal{L}}_{\mathcal{H}} \end{pmatrix}, \qquad \boldsymbol{\mathcal{L}}_{\mathcal{X}}^* = \begin{pmatrix} \boldsymbol{\mathcal{L}}_{\mathcal{V}}^* & 0 \\ 0 & \boldsymbol{\mathcal{L}}_{\mathcal{H}}^* \end{pmatrix}.$$

Finally, we assume throughout this section that $\mathbf{A}$ is a skew-adjoint operator in $\boldsymbol{\mathcal{X}}$, i.e., we have

$$(\mathbf{A}\boldsymbol{\varphi} \mid \boldsymbol{\psi})_{\boldsymbol{\mathcal{X}}} = -(\boldsymbol{\varphi} \mid \mathbf{A}\boldsymbol{\psi})_{\boldsymbol{\mathcal{X}}}, \qquad \boldsymbol{\varphi}, \boldsymbol{\psi} \in \boldsymbol{\mathcal{X}}. \tag{7.57}$$

Again, note that these assumptions are satisfied for the specific examples considered in Chapter 5.

Remark 7.17. *For the ease of presentation, we introduce for $t \in J_T$ and $y = (u, v) \in B_{\mathcal{Y}}(R)$ the equivalent notation*

$$\Lambda_{\mathcal{V}}(y) = \Lambda_{\mathcal{V}}(u,v), \quad \Lambda_{\mathcal{H}}(y) = \Lambda_{\mathcal{H}}(u,v), \quad \mathcal{A}_{\mathcal{V}}(y) = \mathcal{A}_{\mathcal{V}}(u,v), \quad \mathcal{A}_{\mathcal{H}}(y) = \mathcal{A}_{\mathcal{H}}(u,v),$$
$$\mathrm{F}_{\mathcal{V}}(t,y) = \mathrm{F}_{\mathcal{V}}(t,u,v), \quad \mathrm{F}_{\mathcal{H}}(t,y) = \mathrm{F}_{\mathcal{H}}(t,u,v), \quad \mathcal{F}_{\mathcal{V}}(t,y) = \mathcal{F}_{\mathcal{V}}(t,u,v), \quad \mathcal{F}_{\mathcal{H}}(t,y) = \mathcal{F}_{\mathcal{H}}(t,u,v).$$

For the discrete counterparts of these mappings, we proceed analogously.

In [Sturm, 2017, Chap. 4], the application of the leapfrog scheme to the linear Maxwell equations is investigated, based on the interpretation of the leapfrog scheme as a perturbed version of the Crank–Nicolson scheme. We follow a similar approach to analyze the leapfrog scheme applied to quasilinear wave-type equations. However, we interpret the leapfrog scheme as a perturbation of the linearly implicit midpoint rule instead, as we have already established the wellposedness of this scheme. Thus, we first prove an error estimate similar to (7.15) for the full discretization with the linearly implicit midpoint rule.

Therefore, we define

$$\boldsymbol{\mathcal{R}}_-(\boldsymbol{\xi}) \coloneqq \mathrm{Id} - \tfrac{\tau}{2}\boldsymbol{\mathcal{A}}(\boldsymbol{\xi}), \qquad \boldsymbol{\mathcal{R}}_+(\boldsymbol{\xi}) \coloneqq \mathrm{Id} + \tfrac{\tau}{2}\boldsymbol{\mathcal{A}}(\boldsymbol{\xi}), \qquad \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}), \tag{7.58}$$

such that the linearly implicit midpoint rule (7.5) can be written as

$$\boldsymbol{\mathcal{R}}_-(\underline{\boldsymbol{y}}_{n+1/2})\boldsymbol{y}_{n+1} = \boldsymbol{\mathcal{R}}_+(\underline{\boldsymbol{y}}_{n+1/2})\boldsymbol{y}_n + \tau\underline{\boldsymbol{\mathcal{F}}}_{n+1/2}, \qquad n = 0, \dots, N-1. \tag{7.59}$$

The following lemma, which is a direct consequence of [Sturm, 2017, Lem. 4.10], states important properties of these operators.

Lemma 7.18. *Let $\boldsymbol{\varphi}, \boldsymbol{\psi} \in \boldsymbol{\mathcal{X}}$, $\boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$, and (7.57) be satisfied. Then, the nonlinear operators $\boldsymbol{\mathcal{R}}_-(\boldsymbol{\xi})$ and $\boldsymbol{\mathcal{R}}_+(\boldsymbol{\xi})$ have the following properties.*

$$(\boldsymbol{\mathcal{R}}_-(\boldsymbol{\xi})\boldsymbol{\varphi} \mid \boldsymbol{\psi})_{\boldsymbol{\Lambda}(\boldsymbol{\xi})} = (\boldsymbol{\varphi} \mid \boldsymbol{\mathcal{R}}_+(\boldsymbol{\xi})\boldsymbol{\psi})_{\boldsymbol{\Lambda}(\boldsymbol{\xi})}, \tag{7.60a}$$
$$(\boldsymbol{\mathcal{R}}_-(\boldsymbol{\xi})\boldsymbol{\varphi} \mid \boldsymbol{\varphi})_{\boldsymbol{\Lambda}(\boldsymbol{\xi})} = (\boldsymbol{\mathcal{R}}_+(\boldsymbol{\xi})\boldsymbol{\varphi} \mid \boldsymbol{\varphi})_{\boldsymbol{\Lambda}(\boldsymbol{\xi})} = \|\boldsymbol{\varphi}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{\xi})}, \tag{7.60b}$$
$$\|\boldsymbol{\mathcal{R}}_-(\boldsymbol{\xi})^{-1}\boldsymbol{\varphi}\|_{\boldsymbol{\Lambda}(\boldsymbol{\xi})} \leq \|\boldsymbol{\varphi}\|_{\boldsymbol{\Lambda}(\boldsymbol{\xi})}. \tag{7.60c}$$

Based on this lemma, we define for $\boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$ the operator $\boldsymbol{\mathcal{R}}(\boldsymbol{\xi}) \coloneqq \boldsymbol{\mathcal{R}}_-(\boldsymbol{\xi})^{-1}\boldsymbol{\mathcal{R}}_+(\boldsymbol{\xi})$. Furthermore, as we also employ compositions of such operators, we set for $i, j \in \mathbb{N}$ with $j > i$

$$\prod_{k=i}^{j} \boldsymbol{\mathcal{R}}(\boldsymbol{\xi}_k) \coloneqq \boldsymbol{\mathcal{R}}(\boldsymbol{\xi}_j) \cdots \boldsymbol{\mathcal{R}}(\boldsymbol{\xi}_i), \qquad \boldsymbol{\xi}_i, \dots, \boldsymbol{\xi}_j \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}),$$

as well as the special cases

$$\prod_{k=i}^{i} \boldsymbol{\mathcal{R}}(\boldsymbol{\xi}_k) \coloneqq \boldsymbol{\mathcal{R}}(\boldsymbol{\xi}_i), \qquad \prod_{k=i}^{i-1} \boldsymbol{\mathcal{R}}(\boldsymbol{\xi}_k) \coloneqq \mathrm{Id}, \qquad \boldsymbol{\xi}_i \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}).$$

Note that these compositions are ordered.

Motivated by [Sturm, 2017, Lem. 4.11], we now prove that compositions of these operators are bounded.

Lemma 7.19. *Let $\boldsymbol{R}^{\partial_t} > 0$ and*

$$\boldsymbol{z} \in C^1(J_T, B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}^{\partial_t})) \cap C(J_T, B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})).$$

Furthermore, let $K \in \mathbb{N}$ and $s_1, \ldots, s_K \in J_T$, with $0 \le s_1 < \ldots < s_K \le T$ and $\boldsymbol{z}_k = \boldsymbol{z}(s_k)$, for $k = 1, \ldots, K$. Then, we have for $\boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}}$ the bound

$$\left\| \Big(\prod_{k=1}^{K} \boldsymbol{\mathcal{R}}(\boldsymbol{z}_k) \Big) \boldsymbol{\varphi} \right\|_{\boldsymbol{\Lambda}(\boldsymbol{z}_K)} \le \mathrm{e}^{C'(s_K - s_1)} \|\boldsymbol{\varphi}\|_{\boldsymbol{\Lambda}(\boldsymbol{z}_1)}. \tag{7.61}$$

Proof. For $k \le K$ and $\boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}}$ arbitrary, we have due to (7.60a) and (7.60b)

$$\begin{aligned} \|\boldsymbol{\mathcal{R}}(\boldsymbol{z}_k)\boldsymbol{\varphi}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{z}_k)} &= \left(\boldsymbol{\mathcal{R}}_-(\boldsymbol{z}_k)\boldsymbol{\mathcal{R}}(\boldsymbol{z}_k)\boldsymbol{\varphi} \mid \boldsymbol{\mathcal{R}}(\boldsymbol{z}_k)\boldsymbol{\varphi}\right)_{\boldsymbol{\Lambda}(\boldsymbol{z}_k)} \\ &= \left(\boldsymbol{\mathcal{R}}_+(\boldsymbol{z}_k)\boldsymbol{\varphi} \mid \boldsymbol{\mathcal{R}}(\boldsymbol{z}_k)\boldsymbol{\varphi}\right)_{\boldsymbol{\Lambda}(\boldsymbol{z}_k)} \\ &= \left(\boldsymbol{\varphi} \mid \boldsymbol{\mathcal{R}}_+(\boldsymbol{z}_k)\boldsymbol{\varphi}\right)_{\boldsymbol{\Lambda}(\boldsymbol{z}_k)} \\ &= \|\boldsymbol{\varphi}\|^2_{\boldsymbol{\Lambda}(\boldsymbol{z}_k)}. \end{aligned}$$

Hence, $\boldsymbol{\mathcal{R}}(\boldsymbol{z}_k)$ is an isometry on $\boldsymbol{\mathcal{X}}$ with respect to $\|\cdot\|_{\boldsymbol{\Lambda}(\boldsymbol{z}_k)}$. Further, (4.11) implies for $k < K$

$$\|\boldsymbol{\varphi}\|_{\boldsymbol{\Lambda}(\boldsymbol{z}_{k+1})} \le \mathrm{e}^{C'(s_{k+1} - s_k)} \|\boldsymbol{\varphi}\|_{\boldsymbol{\Lambda}(\boldsymbol{z}_k)}.$$

Using these results alternately concludes the proof. □

Note that, contrary to the linear setting, these properties are not sufficient to prove stability of the discrete scheme, as we lack the continuous dependency of the state-dependent norm on time. Thus, we directly consider the corresponding scheme for the error in order to exploit the regularity of the continuous solution.

To do so, we rewrite (7.10) using the operators $\boldsymbol{\mathcal{R}}_-, \boldsymbol{\mathcal{R}}_+$ defined in (7.58). For $n = 0, \ldots, N-1$, this yields the error recursion

$$\boldsymbol{\mathcal{R}}_-(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})\boldsymbol{e}_{n+1} = \boldsymbol{\mathcal{R}}_+(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})\boldsymbol{e}_n + \tau \boldsymbol{g}^{\mathrm{LI}}_{n+1}$$

or equivalently

$$\boldsymbol{e}_{n+1} = \boldsymbol{\mathcal{R}}(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})\boldsymbol{e}_n + \tau \boldsymbol{\mathcal{R}}_-(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})^{-1} \boldsymbol{g}^{\mathrm{LI}}_{n+1},$$

where the right-hand side is again given by (7.12). Using this relation recursively implies for $n = 0, \ldots, N$

$$\boldsymbol{e}_n = \Big(\prod_{k=0}^{n-1} \boldsymbol{\mathcal{R}}(\boldsymbol{\mathcal{I}}\widetilde{y}_{k+1/2}) \Big) \boldsymbol{e}_0 + \tau \sum_{j=0}^{n-1} \Big(\prod_{k=j+1}^{n-1} \boldsymbol{\mathcal{R}}(\boldsymbol{\mathcal{I}}\widetilde{y}_{k+1/2}) \Big) \boldsymbol{\mathcal{R}}_-(\boldsymbol{\mathcal{I}}\widetilde{y}_{j+1/2})^{-1} \boldsymbol{g}^{\mathrm{LI}}_{j+1}. \tag{7.62}$$

For the first term, we obtain with (4.11) and (7.61) the bound

$$\begin{aligned} \left\| \Big(\prod_{k=0}^{n-1} \boldsymbol{\mathcal{R}}(\boldsymbol{\mathcal{I}}\widetilde{y}_{k+1/2}) \Big) \boldsymbol{e}_0 \right\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_n)} &\le \mathrm{e}^{C'(t_n - t_{n-1/2})} \left\| \Big(\prod_{k=0}^{n-1} \boldsymbol{\mathcal{R}}(\boldsymbol{\mathcal{I}}\widetilde{y}_{k+1/2}) \Big) \boldsymbol{e}_0 \right\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{n-1/2})} \\ &\le \mathrm{e}^{C'(t_n - t_{1/2})} \|\boldsymbol{e}_0\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{1/2})} \\ &\le \mathrm{e}^{C' t_n} \|\boldsymbol{e}_0\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_0)}. \end{aligned} \tag{7.63}$$

Similarly, for the second term we have with (4.11), (7.61), and (7.60c) for $j = 0, \dots, n-2$ the estimate

$$
\begin{aligned}
&\Big\| \Big(\prod_{k=j+1}^{n-1} \mathcal{R}(\mathcal{I}\widetilde{y}_{k+1/2}) \Big) \mathcal{R}_-(\mathcal{I}\widetilde{y}_{j+1/2})^{-1} \boldsymbol{g}_{j+1}^{\mathrm{LI}} \Big\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_n)} \\
&\quad \le \mathrm{e}^{C'(t_n - t_{n-1/2})} \Big\| \Big(\prod_{k=j+1}^{n-1} \mathcal{R}(\mathcal{I}\widetilde{y}_{k+1/2}) \Big) \mathcal{R}_-(\mathcal{I}\widetilde{y}_{j+1/2})^{-1} \boldsymbol{g}_{j+1}^{\mathrm{LI}} \Big\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{n-1/2})} \\
&\quad \le \mathrm{e}^{C'(t_n - t_{j+3/2})} \| \mathcal{R}_-(\mathcal{I}\widetilde{y}_{j+1/2})^{-1} \boldsymbol{g}_{j+1}^{\mathrm{LI}} \|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{j+3/2})} \\
&\quad \le \mathrm{e}^{C'(t_n - t_{j+1/2})} \| \mathcal{R}_-(\mathcal{I}\widetilde{y}_{j+1/2})^{-1} \boldsymbol{g}_{j+1}^{\mathrm{LI}} \|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{j+1/2})} \\
&\quad \le \mathrm{e}^{C'(t_n - t_{j+1/2})} \| \boldsymbol{g}_{j+1}^{\mathrm{LI}} \|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{j+1/2})}.
\end{aligned}
\tag{7.64}
$$

Finally, the same argumentation implies for the second term with $j = n-1$

$$
\begin{aligned}
\Big\| \Big(\prod_{k=n}^{n-1} \mathcal{R}(\mathcal{I}\widetilde{y}_{k+1/2}) \Big) \mathcal{R}_-(\mathcal{I}\widetilde{y}_{n-1/2})^{-1} \boldsymbol{g}_{n}^{\mathrm{LI}} \Big\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_n)} &= \| \mathcal{R}_-(\mathcal{I}\widetilde{y}_{n-1/2})^{-1} \boldsymbol{g}_{n}^{\mathrm{LI}} \|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_n)} \\
&\le \mathrm{e}^{C'(t_n - t_{n-1/2})} \| \mathcal{R}_-(\mathcal{I}\widetilde{y}_{n-1/2})^{-1} \boldsymbol{g}_{n}^{\mathrm{LI}} \|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{n-1/2})} \\
&\le \mathrm{e}^{C'(t_n - t_{n-1/2})} \| \boldsymbol{g}_{n}^{\mathrm{LI}} \|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{n-1/2})}.
\end{aligned}
\tag{7.65}
$$

Thus, using these bounds in (7.62) proves

$$
\mathrm{e}^{-C't_n} \| \boldsymbol{e}_n \|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_n)} \le \| \boldsymbol{e}_0 \|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_0)} + \tau \sum_{j=0}^{n-1} \mathrm{e}^{-C't_{j+1/2}} \| \boldsymbol{g}_{j+1}^{\mathrm{LI}} \|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{j+1/2})}.
\tag{7.66}
$$

As in Section 7.1.1, we have to take special care of the first step $n = 0$ in order to get second order convergence. Hence, we derive from (7.34), (7.36), (4.11), and Lemma 7.6 the bound

$$
\begin{aligned}
\| \boldsymbol{g}_1^{\mathrm{LI}} \|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{1/2})} &\le C \big(\mathrm{e}^{C't_{1/2}} \| \boldsymbol{e}_0 \|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_0)} + \tau^2 \sup_{[t_0,t_1]} \| \partial_t^3 y \|_{\mathcal{X}} + \tau^2 \sup_{[t_0,t_1]} \| \partial_t^2 y \|_{\mathcal{Y}} + \tau \sup_{[t_0,t_{1/2}]} \| \partial_t y \|_{\mathcal{Y}} \\
&\quad + \| (\mathcal{I} - \mathcal{J}) \widetilde{y}_{1/2} \|_{\boldsymbol{\mathcal{X}}} + \sup_{[t_0,t_1]} \| \mathcal{R}_\Lambda(\widetilde{y}_{1/2}) \partial_t y \|_{\boldsymbol{\mathcal{X}}} + \| \mathcal{R}_{\mathrm{A}}(\widetilde{y}_1 + \widetilde{y}_0) \|_{\boldsymbol{\mathcal{X}}} + \| \mathcal{R}_{\mathrm{F}}(t_{1/2}, \widetilde{y}_{1/2}) \|_{\boldsymbol{\mathcal{X}}} \big).
\end{aligned}
\tag{7.67}
$$

Similarly, we derive from (7.34), (7.39), (4.11), and Lemma 7.6 for $j > 0$

$$
\begin{aligned}
\| \boldsymbol{g}_{j+1}^{\mathrm{LI}} \|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{j+1/2})} &\le C \big(\mathrm{e}^{C'(t_{j+1/2} - t_j)} \| \boldsymbol{e}_j \|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_j)} + \mathrm{e}^{C'(t_{j+1/2} - t_{j-1})} \| \boldsymbol{e}_{j-1} \|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{j-1})} \\
&\quad + \tau^2 \big(\sup_{[t_j,t_{j+1}]} \| \partial_t^3 y \|_{\mathcal{X}} + \sup_{[t_{j-1},t_{j+1}]} \| \partial_t^2 y \|_{\mathcal{Y}} \big) + \| (\mathcal{I} - \mathcal{J}) \widetilde{y}_{j+1/2} \|_{\boldsymbol{\mathcal{X}}} \\
&\quad + \sup_{[t_j,t_{j+1}]} \| \mathcal{R}_\Lambda(\widetilde{y}_{j+1/2}) \partial_t y \|_{\boldsymbol{\mathcal{X}}} + \| \mathcal{R}_{\mathrm{A}}(\widetilde{y}_{j+1} + \widetilde{y}_j) \|_{\boldsymbol{\mathcal{X}}} + \| \mathcal{R}_{\mathrm{F}}(t_{j+1/2}, \widetilde{y}_{j+1/2}) \|_{\boldsymbol{\mathcal{X}}} \big).
\end{aligned}
\tag{7.68}
$$

Since these bounds imply together with (7.66)

$$
\begin{aligned}
\mathrm{e}^{-C't_n} \| \boldsymbol{e}_n \|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_n)} &\le C \| \boldsymbol{e}_0 \|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_0)} + C\tau \sum_{j=0}^{n-1} \mathrm{e}^{-C't_j} \| \boldsymbol{e}_j \|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_j)} \\
&\quad + Ct_n \Big(\sup_{[0,t_n]} \| (\mathcal{I} - \mathcal{J}) y \|_{\boldsymbol{\mathcal{X}}} + \tau^2 \big(\sup_{[0,t_n]} \| \partial_t y \|_{\mathcal{Y}} + \sup_{[0,t_n]} \| \partial_t^2 y \|_{\mathcal{Y}} + \sup_{[0,t_n]} \| \partial_t^3 y \|_{\mathcal{X}} \big) \\
&\quad + \sup_{s_1,s_2 \in [0,t_n]} \| \mathcal{R}_\Lambda(y(s_1)) \partial_t y(s_2) \|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_n]} \| \mathcal{R}_{\mathrm{A}} y \|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_n]} \| \mathcal{R}_{\mathrm{F}}(\cdot, y) \|_{\boldsymbol{\mathcal{X}}} \Big),
\end{aligned}
$$

the discrete Gronwall inequality finally proves

$$\begin{aligned}\|\boldsymbol{e}_n\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_n)} \le C(1+t_n)\mathrm{e}^{Ct_n}\Big(&\|\boldsymbol{e}_0\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_0)} + \sup_{[0,t_n]}\|(\boldsymbol{\mathcal{I}}-\boldsymbol{\mathcal{J}})y\|_{\boldsymbol{\mathcal{X}}}\\ &+\tau^2\big(\sup_{[0,t_n]}\|\partial_t^2 y\|_{\mathcal{Y}} + \sup_{[0,t_n]}\|\partial_t^2 y\|_{\mathcal{Y}} + \sup_{[0,t_n]}\|\partial_t^3 y\|_{\mathcal{X}}\big)\\ &+\sup_{s_1,s_2\in[0,t_n]}\|\mathcal{R}_\Lambda(y(s_1))\partial_t y(s_2)\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_n]}\|\mathcal{R}_{\mathrm{A}}y\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_n]}\|\mathcal{R}_{\mathrm{F}}(\cdot,y)\|_{\boldsymbol{\mathcal{X}}}\Big),\end{aligned}$$

with a constant $C>0$ independent of τ and h. Note that this corresponds to (7.15).

To investigate the leapfrog scheme, we first elaborate on the connection of this scheme to the linearly implicit midpoint rule. We then prove wellposedness and an error estimate for the leapfrog scheme based on the results from the previous subsection.

Note that throughout this subsection, we employ the additional assumptions

$$\Lambda_{\mathcal{V}}(\xi) \equiv \mathrm{Id}, \qquad\qquad \boldsymbol{\Lambda}_{\boldsymbol{\mathcal{V}}}(\boldsymbol{\xi}) \equiv \mathrm{Id}, \tag{7.69}$$

for $\xi \in B_{\mathcal{Y}}(R)$ and $\boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$. For a discussion on this assumption, see Remark 7.28.

As the leapfrog scheme is an explicit scheme for problems of the form (6.8) and hence computationally cheap, it is a very appealing scheme. However, note that a direct application of the leapfrog scheme to (7.56) yields an implicit scheme, as the right-hand sides depend on both unknowns. Hence, we consider in the following a variant of the leapfrog scheme, which is based on the linearly implicit midpoint rule (7.5), but nevertheless explicit. The basic idea is to employ the approximation (7.6) for the intermediate values. More precisely, we consider the following modified version of the leapfrog scheme applied to (7.56), which is for $n=0,\dots,N-1$ and $\boldsymbol{y}_n = (\boldsymbol{u}_n,\boldsymbol{v}_n)$ given by

$$\widehat{\boldsymbol{u}}_{n+1/2} = \boldsymbol{u}_n + \tfrac{\tau}{2}\mathbf{A}_{\boldsymbol{\mathcal{V}}}\boldsymbol{v}_n + \tfrac{\tau}{2}\boldsymbol{\mathcal{F}}_{\boldsymbol{\mathcal{V}}}(t_{n+1/2},\underline{\boldsymbol{y}}_{n+1/2}), \tag{7.70a}$$

$$\boldsymbol{v}_{n+1} = \boldsymbol{v}_n + \tau\boldsymbol{\mathcal{A}}_{\boldsymbol{\mathcal{H}}}(\underline{\boldsymbol{y}}_{n+1/2})\widehat{\boldsymbol{u}}_{n+1/2} + \tau\boldsymbol{\mathcal{F}}_{\boldsymbol{\mathcal{H}}}(t_{n+1/2},\underline{\boldsymbol{y}}_{n+1/2}), \tag{7.70b}$$

$$\boldsymbol{u}_{n+1} = \widehat{\boldsymbol{u}}_{n+1/2} + \tfrac{\tau}{2}\mathbf{A}_{\boldsymbol{\mathcal{V}}}\boldsymbol{v}_{n+1} + \tfrac{\tau}{2}\boldsymbol{\mathcal{F}}_{\boldsymbol{\mathcal{V}}}(t_{n+1/2},\underline{\boldsymbol{y}}_{n+1/2}). \tag{7.70c}$$

In order to elaborate on the correlation to the linearly implicit midpoint rule, we use (7.70a) in (7.70c) to eliminate $\widehat{\boldsymbol{u}}_{n+1/2}$. Furthermore, (7.70a) and (7.70c) imply

$$\widehat{\boldsymbol{u}}_{n+1/2} = \frac{\boldsymbol{u}_{n+1}+\boldsymbol{u}_n}{2} + \tfrac{\tau}{4}\mathbf{A}_{\boldsymbol{\mathcal{V}}}(\boldsymbol{v}_n - \boldsymbol{v}_{n+1}).$$

Using this relation to eliminate $\widehat{\boldsymbol{u}}_{n+1/2}$ in (7.70b), we hence obtain that, for $n=0,\dots,N-1$,

$$\begin{aligned}\boldsymbol{u}_{n+1} &= \boldsymbol{u}_n + \tau\mathbf{A}_{\boldsymbol{\mathcal{V}}}\frac{\boldsymbol{v}_{n+1}+\boldsymbol{v}_n}{2} + \tau\boldsymbol{\mathcal{F}}_{\boldsymbol{\mathcal{V}}}(t_{n+1/2},\underline{\boldsymbol{y}}_{n+1/2}),\\ \boldsymbol{v}_{n+1} &= \boldsymbol{v}_n + \tau\boldsymbol{\mathcal{A}}_{\boldsymbol{\mathcal{H}}}(\underline{\boldsymbol{y}}_{n+1/2})\frac{\boldsymbol{u}_{n+1}+\boldsymbol{u}_n}{2} + \tau\boldsymbol{\mathcal{F}}_{\boldsymbol{\mathcal{H}}}(t_{n+1/2},\underline{\boldsymbol{y}}_{n+1/2}) + \tfrac{\tau^2}{4}\boldsymbol{\mathcal{A}}_{\boldsymbol{\mathcal{H}}}(\underline{\boldsymbol{y}}_{n+1/2})\mathbf{A}_{\boldsymbol{\mathcal{V}}}(\boldsymbol{v}_n - \boldsymbol{v}_{n+1})\end{aligned}$$

is an equivalent formulation of the leapfrog scheme. Thus, similarly to the linearly implicit midpoint rule in (7.59), we can rewrite (7.70) as

$$\widehat{\boldsymbol{\mathcal{R}}}_-(\underline{\boldsymbol{y}}_{n+1/2})\boldsymbol{y}_{n+1} = \widehat{\boldsymbol{\mathcal{R}}}_+(\underline{\boldsymbol{y}}_{n+1/2})\boldsymbol{y}_n + \tau\underline{\boldsymbol{\mathcal{F}}}_{n+1/2}, \qquad n=0,\dots,N-1, \tag{7.71}$$

for nonlinear operators $\widehat{\mathfrak{R}}_-, \widehat{\mathfrak{R}}_+, \mathfrak{D}$ given by

$$\mathfrak{D}(\boldsymbol{\xi}) := \begin{pmatrix} 0 & 0 \\ 0 & \frac{\tau}{2}\boldsymbol{\mathcal{A}}_{\mathcal{H}}(\boldsymbol{\xi})\mathbf{A}_{\mathcal{V}} \end{pmatrix}, \qquad \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}), \tag{7.72}$$

and

$$\widehat{\mathfrak{R}}_-(\boldsymbol{\xi}) := \mathfrak{R}_-(\boldsymbol{\xi}) + \tfrac{\tau}{2}\mathfrak{D}(\boldsymbol{\xi}), \qquad \widehat{\mathfrak{R}}_+(\boldsymbol{\xi}) := \mathfrak{R}_+(\boldsymbol{\xi}) + \tfrac{\tau}{2}\mathfrak{D}(\boldsymbol{\xi}), \qquad \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}). \tag{7.73}$$

For the sake of presentation, we also use the short notation from (7.1), (7.2), (7.13), and (7.43) for $\mathfrak{D}$, i.e., we write

$$\widetilde{\mathfrak{D}}_n := \mathfrak{D}(\boldsymbol{\mathcal{I}}\widetilde{y}_n), \qquad \underline{\mathfrak{D}}_n := \mathfrak{D}(\underline{\boldsymbol{y}}_n), \tag{7.74}$$

In the following lemma, we state important properties of the operators from (7.72) and (7.73) based on [Sturm, 2017, Lem. 4.10, 4.14, and 4.15]. For $\vartheta \in (0,1)$ arbitrary but fixed, this is based on the classical CFL condition of the leapfrog scheme

$$\tau \boldsymbol{C}_{\mathbf{A}}(h) \leq 2\vartheta \boldsymbol{c}_{\mathbf{\Lambda}}, \tag{7.75}$$

with constant $\boldsymbol{C}_{\mathbf{A}}(h)$ as given in (4.29).

Lemma 7.20. *Let $\boldsymbol{\varphi}, \boldsymbol{\psi} \in \boldsymbol{\mathcal{X}}$, $\boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$, and $\mathbf{A}$ skew-adjoint in $\boldsymbol{\mathcal{X}}$. Further, let $\vartheta \in (0,1)$ arbitrary but fixed. If (7.69) is satisfied, we have*

$$\left(\widehat{\mathfrak{R}}_-(\boldsymbol{\xi})\boldsymbol{\varphi} \mid \boldsymbol{\psi}\right)_{\mathbf{\Lambda}(\boldsymbol{\xi})} = \left(\boldsymbol{\varphi} \mid \widehat{\mathfrak{R}}_+(\boldsymbol{\xi})\boldsymbol{\psi}\right)_{\mathbf{\Lambda}(\boldsymbol{\xi})}, \tag{7.76a}$$

$$\left(\widehat{\mathfrak{R}}_-(\boldsymbol{\xi})\boldsymbol{\varphi} \mid \boldsymbol{\varphi}\right)_{\mathbf{\Lambda}(\boldsymbol{\xi})} = \left(\widehat{\mathfrak{R}}_+(\boldsymbol{\xi})\boldsymbol{\varphi} \mid \boldsymbol{\varphi}\right)_{\mathbf{\Lambda}(\boldsymbol{\xi})} = \|\boldsymbol{\varphi}\|^2_{\mathbf{\Lambda}(\boldsymbol{\xi})} - \tfrac{\tau^2}{4}\|\mathbf{A}_{\mathcal{V}}\boldsymbol{\varphi}\|^2_{\boldsymbol{\mathcal{X}}_{\mathcal{V}}}. \tag{7.76b}$$

If additionally (7.75) is satisfied, we have for $C_{\mathrm{stb}} = (1-\vartheta^2)^{-1}$

$$(1-\vartheta^2)\|\boldsymbol{\varphi}\|^2_{\mathbf{\Lambda}(\boldsymbol{\xi})} \leq \left(\widehat{\mathfrak{R}}_-(\boldsymbol{\xi})\boldsymbol{\varphi} \mid \boldsymbol{\varphi}\right)_{\mathbf{\Lambda}(\boldsymbol{\xi})} \leq \|\boldsymbol{\varphi}\|^2_{\mathbf{\Lambda}(\boldsymbol{\xi})}, \tag{7.76c}$$

as well as the bound

$$\|\widehat{\mathfrak{R}}_-(\boldsymbol{\xi})^{-1}\boldsymbol{\varphi}\|_{\mathbf{\Lambda}(\boldsymbol{\xi})} \leq C_{\mathrm{stb}}\|\boldsymbol{\varphi}\|_{\mathbf{\Lambda}(\boldsymbol{\xi})}. \tag{7.76d}$$

Proof. Let $\boldsymbol{\varphi} = (\boldsymbol{\varphi}_{\mathcal{V}}, \boldsymbol{\varphi}_{\mathcal{H}}), \boldsymbol{\psi} = (\boldsymbol{\psi}_{\mathcal{V}}, \boldsymbol{\psi}_{\mathcal{H}}) \in \boldsymbol{\mathcal{X}}$ and $\boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$. Due to the skew-adjointness (7.57) of $\mathbf{A}$, we obtain the equation

$$(\mathbf{A}_{\mathcal{V}}\boldsymbol{\varphi}_{\mathcal{H}} \mid \boldsymbol{\psi}_{\mathcal{V}})_{\boldsymbol{\mathcal{X}}_{\mathcal{V}}} + (\mathbf{A}_{\mathcal{H}}\boldsymbol{\varphi}_{\mathcal{V}} \mid \boldsymbol{\psi}_{\mathcal{H}})_{\boldsymbol{\mathcal{X}}_{\mathcal{H}}} = -(\boldsymbol{\varphi}_{\mathcal{V}} \mid \mathbf{A}_{\mathcal{V}}\boldsymbol{\psi}_{\mathcal{H}})_{\boldsymbol{\mathcal{X}}_{\mathcal{V}}} - (\boldsymbol{\varphi}_{\mathcal{H}} \mid \mathbf{A}_{\mathcal{H}}\boldsymbol{\psi}_{\mathcal{V}})_{\boldsymbol{\mathcal{X}}_{\mathcal{H}}}.$$

As this is true for $\boldsymbol{\varphi}, \boldsymbol{\psi}$ arbitrary, we particularly have

$$(\mathbf{A}_{\mathcal{H}}\boldsymbol{\varphi}_{\mathcal{V}} \mid \boldsymbol{\psi}_{\mathcal{H}})_{\boldsymbol{\mathcal{X}}_{\mathcal{H}}} = -(\boldsymbol{\varphi}_{\mathcal{V}} \mid \mathbf{A}_{\mathcal{V}}\boldsymbol{\psi}_{\mathcal{H}})_{\boldsymbol{\mathcal{X}}_{\mathcal{V}}}.$$

Based on this relation, we deduce

$$\begin{aligned}
(\mathfrak{D}(\boldsymbol{\xi})\boldsymbol{\varphi} \mid \boldsymbol{\psi})_{\mathbf{\Lambda}(\boldsymbol{\xi})} &= \tfrac{\tau}{2}(\boldsymbol{\mathcal{A}}_{\mathcal{H}}(\boldsymbol{\xi})\mathbf{A}_{\mathcal{V}}\boldsymbol{\varphi}_{\mathcal{H}} \mid \boldsymbol{\psi}_{\mathcal{H}})_{\mathbf{\Lambda}_{\mathcal{H}}(\boldsymbol{\xi})} \\
&= \tfrac{\tau}{2}(\mathbf{A}_{\mathcal{H}}\mathbf{A}_{\mathcal{V}}\boldsymbol{\varphi}_{\mathcal{H}} \mid \boldsymbol{\psi}_{\mathcal{H}})_{\boldsymbol{\mathcal{X}}_{\mathcal{H}}} \\
&= -\tfrac{\tau}{2}(\mathbf{A}_{\mathcal{V}}\boldsymbol{\varphi}_{\mathcal{H}} \mid \mathbf{A}_{\mathcal{V}}\boldsymbol{\psi}_{\mathcal{H}})_{\boldsymbol{\mathcal{X}}_{\mathcal{V}}}.
\end{aligned}$$

Hence, (7.60a) and (7.60b) yield (7.76a) and (7.76b), respectively. Further, the boundedness (4.29) of $\mathbf{A}$ implies

$$\|\mathbf{A}_{\mathcal{V}}\|_{\mathcal{L}(\mathcal{X}_{\mathcal{H}},\mathcal{X}_{\mathcal{V}})}, \|\mathbf{A}_{\mathcal{H}}\|_{\mathcal{L}(\mathcal{X}_{\mathcal{V}},\mathcal{X}_{\mathcal{H}})} \leq \boldsymbol{C}_{\mathbf{A}}(h). \tag{7.77}$$

Thus, due to the CFL condition (7.75), (7.76c) is a direct consequence of (7.76b). For the last statement, we conclude from (7.76c)

$$\|\widehat{\mathfrak{R}}_-(\boldsymbol{\xi})\varphi\|_{\mathbf{\Lambda}(\boldsymbol{\xi})} = \sup_{\psi\in\mathcal{X}} \frac{\left(\widehat{\mathfrak{R}}_-(\boldsymbol{\xi})\varphi \mid \psi\right)_{\mathbf{\Lambda}(\boldsymbol{\xi})}}{\|\psi\|_{\mathbf{\Lambda}(\boldsymbol{\xi})}} \geq \frac{\left(\widehat{\mathfrak{R}}_-(\boldsymbol{\xi})\varphi \mid \varphi\right)_{\mathbf{\Lambda}(\boldsymbol{\xi})}}{\|\varphi\|_{\mathbf{\Lambda}(\boldsymbol{\xi})}} \geq (1-\vartheta^2)\|\varphi\|_{\mathbf{\Lambda}(\boldsymbol{\xi})}.$$

Hence, $\widehat{\mathfrak{R}}_-(\boldsymbol{\xi})$ is invertible and the choice $\varphi = \widehat{\mathfrak{R}}_-(\boldsymbol{\xi})^{-1}\psi$ finally yields (7.76d). □

As in the analysis of the linearly implicit midpoint rule in the previous section, we now define the operator $\widehat{\mathfrak{R}}(\boldsymbol{\xi}) \coloneqq \widehat{\mathfrak{R}}_-(\boldsymbol{\xi})^{-1}\widehat{\mathfrak{R}}_+(\boldsymbol{\xi})$, for $\boldsymbol{\xi} \in B_{\mathcal{Y}}(\boldsymbol{R})$. Based on [Sturm, 2017, Lem. 4.15], we again prove that compositions of these operators are bounded.

Lemma 7.21. *Let $\vartheta \in (0,1)$ arbitrary but fixed such that (7.75) is satisfied, and $\boldsymbol{R}^{\partial_t} > 0$ such that*

$$\boldsymbol{z} \in C^1(J_T, B_{\mathcal{Y}}(\boldsymbol{R}^{\partial_t})) \cap C(J_T, B_{\mathcal{Y}}(\boldsymbol{R}))$$

holds. Furthermore, for $K \in \mathbb{N}$ and $s_1,\dots,s_K \in J_T$, let $0 \leq s_1 < \dots < s_K \leq T$ and $\boldsymbol{z}_k = \boldsymbol{z}(s_k)$, for $k = 1,\dots,K$. If (7.69) is satisfied, we have for $\varphi \in \mathcal{X}$ the bound

$$\left\| \Big(\prod_{k=1}^{K} \widehat{\mathfrak{R}}(\boldsymbol{z}_k)\Big)\varphi \right\|_{\mathbf{\Lambda}(\boldsymbol{z}_K)} \leq C_{\mathrm{stb}}^{\frac{1}{2}} \mathrm{e}^{C'(s_K-s_1)} \|\varphi\|_{\mathbf{\Lambda}(\boldsymbol{z}_1)},$$

where the constant C' is given in Lemma 4.2.

Proof. We first observe for $k,\ell \leq K$

$$\mathbf{\Lambda}(\boldsymbol{z}_k)\widehat{\mathfrak{R}}_-(\boldsymbol{z}_k) - \mathbf{\Lambda}(\boldsymbol{z}_\ell)\widehat{\mathfrak{R}}_-(\boldsymbol{z}_\ell) = \mathbf{\Lambda}(\boldsymbol{z}_k) - \mathbf{\Lambda}(\boldsymbol{z}_\ell), \tag{7.78}$$

which yields for $k < K$ and $\varphi \in \mathcal{X}$

$$\begin{aligned}
\left(\widehat{\mathfrak{R}}_-(\boldsymbol{z}_{k+1})\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\varphi \mid \widehat{\mathfrak{R}}(\boldsymbol{z}_k)\varphi\right)_{\mathbf{\Lambda}(\boldsymbol{z}_{k+1})} &= \left(\mathbf{\Lambda}(\boldsymbol{z}_{k+1})\widehat{\mathfrak{R}}_-(\boldsymbol{z}_{k+1})\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\varphi \mid \widehat{\mathfrak{R}}(\boldsymbol{z}_k)\varphi\right)_{\mathcal{X}} \\
&= \left(\widehat{\mathfrak{R}}_-(\boldsymbol{z}_k)\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\varphi \mid \widehat{\mathfrak{R}}(\boldsymbol{z}_k)\varphi\right)_{\mathbf{\Lambda}(\boldsymbol{z}_k)} \\
&\quad + \left((\mathbf{\Lambda}(\boldsymbol{z}_{k+1}) - \mathbf{\Lambda}(\boldsymbol{z}_k))\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\varphi \mid \widehat{\mathfrak{R}}(\boldsymbol{z}_k)\varphi\right)_{\mathcal{X}}.
\end{aligned}$$

Furthermore, the Cauchy–Schwarz inequality and the Lipschitz continuity (4.4a) of $\mathbf{\Lambda}$ yield

$$\begin{aligned}
\left((\mathbf{\Lambda}(\boldsymbol{z}_{k+1}) - \mathbf{\Lambda}(\boldsymbol{z}_k))\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\varphi \mid \widehat{\mathfrak{R}}(\boldsymbol{z}_k)\varphi\right)_{\mathcal{X}} &\leq \|\mathbf{\Lambda}(\boldsymbol{z}_{k+1}) - \mathbf{\Lambda}(\boldsymbol{z}_k)\|_{\mathcal{L}(\mathcal{X})}\|\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\varphi\|_{\mathcal{X}}^2 \\
&\leq \boldsymbol{L}_{\mathbf{\Lambda}}^{\mathcal{X}}\|\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\varphi\|_{\mathcal{X}}^2\|\boldsymbol{z}_{k+1} - \boldsymbol{z}_k\|_{\mathcal{Y}} \\
&\leq \boldsymbol{L}_{\mathbf{\Lambda}}^{\mathcal{X}}\|\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\varphi\|_{\mathcal{X}}^2 \int_{s_k}^{s_{k+1}} \|\partial_t \boldsymbol{z}(s)\|_{\mathcal{Y}}\,\mathrm{d}s \\
&\leq \boldsymbol{L}_{\mathbf{\Lambda}}^{\mathcal{X}}\boldsymbol{R}^{\partial_t}\|\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\varphi\|_{\mathcal{X}}^2(s_{k+1} - s_k).
\end{aligned}$$

The norm equivalence (4.10) and (7.76b) further imply

$$\|\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\boldsymbol{\varphi}\|_{\mathcal{X}}^2 \leq \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1} \left(\widehat{\mathfrak{R}}_-(\boldsymbol{z}_k)\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\boldsymbol{\varphi} \mid \widehat{\mathfrak{R}}(\boldsymbol{z}_k)\boldsymbol{\varphi}\right)_{\boldsymbol{\Lambda}(\boldsymbol{z}_k)}.$$

Hence, since we obtain similarly to the proof of Lemma 7.19 due to (7.76a) and (7.76b)

$$\left(\widehat{\mathfrak{R}}_-(\boldsymbol{z}_k)\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\boldsymbol{\varphi} \mid \widehat{\mathfrak{R}}(\boldsymbol{z}_k)\boldsymbol{\varphi}\right)_{\boldsymbol{\Lambda}(\boldsymbol{z}_k)} = \left(\widehat{\mathfrak{R}}_+(\boldsymbol{z}_k)\boldsymbol{\varphi} \mid \widehat{\mathfrak{R}}(\boldsymbol{z}_k)\boldsymbol{\varphi}\right)_{\boldsymbol{\Lambda}(\boldsymbol{z}_k)} = \left(\widehat{\mathfrak{R}}_-(\boldsymbol{z}_k)\boldsymbol{\varphi} \mid \boldsymbol{\varphi}\right)_{\boldsymbol{\Lambda}(\boldsymbol{z}_k)},$$

we finally have

$$\left(\widehat{\mathfrak{R}}_-(\boldsymbol{z}_{k+1})\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\boldsymbol{\varphi} \mid \widehat{\mathfrak{R}}(\boldsymbol{z}_k)\boldsymbol{\varphi}\right)_{\boldsymbol{\Lambda}(\boldsymbol{z}_{k+1})} \leq \mathrm{e}^{2C'(s_{k+1}-s_k)} \left(\widehat{\mathfrak{R}}_-(\boldsymbol{z}_k)\boldsymbol{\varphi} \mid \boldsymbol{\varphi}\right)_{\boldsymbol{\Lambda}(\boldsymbol{z}_k)},$$

with the constant C' given in Lemma 4.2. Using this result iteratively together with (7.76c) concludes the proof. □

With these preliminary lemmas at hand, we now turn towards the analysis of the full discretization with the leapfrog scheme. As before, the first step is to derive a recursion formula for the discrete error

$$\boldsymbol{e}_n := \boldsymbol{\mathcal{J}}\widetilde{y}_n - \boldsymbol{y}_n. \tag{7.79}$$

For $n = 0, \dots, N-1$, this directly yields

$$\begin{aligned}\widehat{\mathfrak{R}}_-(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})\boldsymbol{e}_{n+1} - \widehat{\mathfrak{R}}_+(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})\boldsymbol{e}_n &= \widehat{\mathfrak{R}}_-(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})\boldsymbol{\mathcal{J}}\widetilde{y}_{n+1} - \widehat{\mathfrak{R}}_+(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})\boldsymbol{\mathcal{J}}\widetilde{y}_n \\ &\quad + \widehat{\mathfrak{R}}_+(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})\boldsymbol{y}_n - \widehat{\mathfrak{R}}_-(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})\boldsymbol{y}_{n+1}.\end{aligned} \tag{7.80}$$

We now consider the terms on the right-hand side separately. For the first term, the definition of $\widehat{\mathfrak{R}}_-$ implies

$$\widehat{\mathfrak{R}}_-(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})\boldsymbol{\mathcal{J}}\widetilde{y}_{n+1} = \boldsymbol{\mathcal{J}}\widetilde{y}_{n+1} - \tfrac{\tau}{2}\widetilde{\boldsymbol{\mathcal{A}}}_{n+1/2}\boldsymbol{\mathcal{J}}\widetilde{y}_{n+1} + \tfrac{\tau}{2}\widetilde{\mathfrak{D}}_{n+1/2}\boldsymbol{\mathcal{J}}\widetilde{y}_{n+1}.$$

Analogously, we obtain from the definition of $\widehat{\mathfrak{R}}_+$ for the second term

$$\widehat{\mathfrak{R}}_+(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})\boldsymbol{\mathcal{J}}\widetilde{y}_n = \boldsymbol{\mathcal{J}}\widetilde{y}_n + \tfrac{\tau}{2}\widetilde{\boldsymbol{\mathcal{A}}}_{n+1/2}\boldsymbol{\mathcal{J}}\widetilde{y}_n + \tfrac{\tau}{2}\widetilde{\mathfrak{D}}_{n+1/2}\boldsymbol{\mathcal{J}}\widetilde{y}_n.$$

Thus, we get with (7.8) for the first difference in (7.80)

$$\begin{aligned}\widehat{\mathfrak{R}}_-(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})\boldsymbol{\mathcal{J}}\widetilde{y}_{n+1} - \widehat{\mathfrak{R}}_+(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})\boldsymbol{\mathcal{J}}\widetilde{y}_n &= \boldsymbol{\mathcal{J}}\big(\widetilde{y}_{n+1} - \widetilde{y}_n\big) - \tau\widetilde{\boldsymbol{\mathcal{A}}}_{n+1/2}\boldsymbol{\mathcal{J}}\big(\widetilde{y}_{n+1/2} - \widehat{\delta}_{n+1/2}\big) \\ &\quad + \tfrac{\tau}{2}\widetilde{\mathfrak{D}}_{n+1/2}\boldsymbol{\mathcal{J}}\big(\widetilde{y}_{n+1} - \widetilde{y}_n\big).\end{aligned}$$

As for the implicit midpoint rules, we now apply the adjoint lift operator $\boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{n+1/2}]$ to (7.7) and add the result to the right-hand side of the previous equation. This proves

$$\begin{aligned}\widehat{\mathfrak{R}}_-(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})\boldsymbol{\mathcal{J}}\widetilde{y}_{n+1} - \widehat{\mathfrak{R}}_+(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})\boldsymbol{\mathcal{J}}\widetilde{y}_n &= \big(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{n+1/2}]\big)\big(\widetilde{y}_{n+1} - \widetilde{y}_n\big) \\ &\quad + \tau\big(\widetilde{\boldsymbol{\mathcal{A}}}_{n+1/2}\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{n+1/2}]\widetilde{\mathcal{A}}_{n+1/2}\big)\big(\widehat{\delta}_{n+1/2} - \widetilde{y}_{n+1/2}\big) \\ &\quad + \tau\boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{n+1/2}]\widetilde{\mathcal{F}}_{n+1/2} + \boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{n+1/2}]\widetilde{\mathcal{A}}_{n+1/2}\widehat{\delta}_{n+1/2} \\ &\quad + \boldsymbol{\mathcal{L}}_\Lambda^*[\widetilde{y}_{n+1/2}]\delta_{n+1} + \tfrac{\tau}{2}\widetilde{\mathfrak{D}}_{n+1/2}\boldsymbol{\mathcal{J}}\big(\widetilde{y}_{n+1} - \widetilde{y}_n\big).\end{aligned} \tag{7.81}$$

For the second difference in (7.80), we obtain again from the definition of $\widehat{\mathfrak{R}}_-$

$$\widehat{\mathfrak{R}}_-(\mathcal{I}\widetilde{y}_{n+1/2})\boldsymbol{y}_{n+1} = \widehat{\mathfrak{R}}_-(\underline{\boldsymbol{y}}_{n+1/2})\boldsymbol{y}_n + \tfrac{\tau}{2}(\underline{\boldsymbol{\mathcal{A}}}_{n+1/2} - \widetilde{\boldsymbol{\mathcal{A}}}_{n+1/2})\boldsymbol{y}_n - \tfrac{\tau}{2}(\underline{\mathfrak{D}}_{n+1/2} - \widetilde{\mathfrak{D}}_{n+1/2})\boldsymbol{y}_n.$$

We further get from the definition of $\widehat{\mathfrak{R}}_+$

$$\widehat{\mathfrak{R}}_+(\mathcal{I}\widetilde{y}_{n+1/2})\boldsymbol{y}_n = \widehat{\mathfrak{R}}_+(\underline{\boldsymbol{y}}_{n+1/2})\boldsymbol{y}_{n+1} - \tfrac{\tau}{2}(\underline{\boldsymbol{\mathcal{A}}}_{n+1/2} - \widetilde{\boldsymbol{\mathcal{A}}}_{n+1/2})\boldsymbol{y}_{n+1} - \tfrac{\tau}{2}(\underline{\mathfrak{D}}_{n+1/2} - \widetilde{\mathfrak{D}}_{n+1/2})\boldsymbol{y}_{n+1}.$$

Hence, the scheme (7.71) implies for the second difference in (7.80)

$$\begin{aligned}\widehat{\mathfrak{R}}_-(\mathcal{I}\widetilde{y}_{n+1/2})\boldsymbol{y}_{n+1} - \widehat{\mathfrak{R}}_+(\mathcal{I}\widetilde{y}_{n+1/2})\boldsymbol{y}_n &= -\tfrac{\tau}{2}(\underline{\mathfrak{D}}_{n+1/2} - \widetilde{\mathfrak{D}}_{n+1/2})(\boldsymbol{y}_{n+1} - \boldsymbol{y}_n) \\ &\quad + \tau\underline{\boldsymbol{\mathcal{F}}}_{n+1/2} + \tau(\underline{\boldsymbol{\mathcal{A}}}_{n+1/2} - \widetilde{\boldsymbol{\mathcal{A}}}_{n+1/2})\boldsymbol{y}_{n+1/2}.\end{aligned} \tag{7.82}$$

Finally, using (7.81) and (7.82) in (7.80), we obtain

$$\widehat{\mathfrak{R}}_-(\mathcal{I}\widetilde{y}_{n+1/2})\boldsymbol{e}_{n+1} = \widehat{\mathfrak{R}}_+(\mathcal{I}\widetilde{y}_{n+1/2})\boldsymbol{e}_n + \tau(\boldsymbol{g}^{\mathrm{LI}}_{n+1} + \boldsymbol{d}_{n+1} + \widehat{\boldsymbol{d}}_{n+1}), \tag{7.83}$$

where $\boldsymbol{g}^{\mathrm{LI}}_{n+1}$ is given by (7.12). The defects $\boldsymbol{d}_{n+1}$ and $\widehat{\boldsymbol{d}}_{n+1}$ are defined by

$$\boldsymbol{d}_{n+1} = \tfrac{1}{2}(\underline{\mathfrak{D}}_{n+1/2} - \widetilde{\mathfrak{D}}_{n+1/2})(\boldsymbol{y}_{n+1} - \boldsymbol{y}_n), \tag{7.84a}$$

$$\widehat{\boldsymbol{d}}_{n+1} = \tfrac{1}{2}\widetilde{\mathfrak{D}}_{n+1/2}\boldsymbol{\mathcal{J}}(\widetilde{y}_{n+1} - \widetilde{y}_n). \tag{7.84b}$$

Using (7.83) recursively implies

$$\boldsymbol{e}_n = \Big(\prod_{k=0}^{n-1}\widehat{\mathfrak{R}}(\mathcal{I}\widetilde{y}_{k+1/2})\Big)\boldsymbol{e}_0 + \tau\sum_{j=0}^{n-1}\Big(\prod_{k=j+1}^{n-1}\widehat{\mathfrak{R}}(\mathcal{I}\widetilde{y}_{k+1/2})\Big)\widehat{\mathfrak{R}}_-(\mathcal{I}\widetilde{y}_{j+1/2})^{-1}(\boldsymbol{g}^{\mathrm{LI}}_{j+1} + \boldsymbol{d}_{j+1} + \widehat{\boldsymbol{d}}_{j+1}).$$

Thus, the same argumentation as in (7.63) to (7.65) with Lemma 7.21 and (7.76d) instead of Lemma 7.19 and (7.60c), respectively, proves for $n = 0, \dots, N$ the estimate

$$\begin{aligned}\mathrm{e}^{-C't_n}\|\boldsymbol{e}_n\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_n)} &\le C_{\mathrm{stb}}^{\frac{1}{2}}\|\boldsymbol{e}_0\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_0)} + C_{\mathrm{stb}}^{\frac{3}{2}}\tau\sum_{j=0}^{n-1}\mathrm{e}^{-C't_{j+1/2}}\|\boldsymbol{g}^{\mathrm{LI}}_{j+1}\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{j+1/2})} \\ &\quad + C_{\mathrm{stb}}^{\frac{1}{2}}\tau\sum_{j=0}^{n-1}\mathrm{e}^{-C't_{j+1/2}}\|\widehat{\mathfrak{R}}_-(\mathcal{I}\widetilde{y}_{j+1/2})^{-1}\boldsymbol{d}_{j+1}\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{j+1/2})} \\ &\quad + \tau\mathrm{e}^{-C't_n}\Big\|\sum_{j=0}^{n-1}\Big(\prod_{k=j+1}^{n-1}\widehat{\mathfrak{R}}(\mathcal{I}\widetilde{y}_{k+1/2})\Big)\widehat{\mathfrak{R}}_-(\mathcal{I}\widetilde{y}_{j+1/2})^{-1}\widehat{\boldsymbol{d}}_{j+1}\Big\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_n)}.\end{aligned} \tag{7.85}$$

This estimate is the basis for the analysis of the full discretization with the leapfrog scheme. Since we already studied $\boldsymbol{g}^{\mathrm{LI}}_{j+1}$ in the analysis of the linearly implicit midpoint rule we only bound $\boldsymbol{d}_{n+1}$ and $\widehat{\boldsymbol{d}}_{n+1}$ in the following. However, note that we have to take special care of $\widehat{\boldsymbol{d}}_{n+1}$ in order to avoid additional assumptions. Thus, we first prove the following lemma for $\boldsymbol{d}_{n+1}$.

Lemma 7.22. *Let $\boldsymbol{\xi}, \boldsymbol{\zeta} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$ and $\boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}}$. If (7.69) is satisfied, we have*

$$\|(\mathfrak{D}(\boldsymbol{\xi}) - \mathfrak{D}(\boldsymbol{\zeta}))\boldsymbol{\varphi}\|_{\boldsymbol{\Lambda}(\boldsymbol{\xi})} \le C\|\mathfrak{D}(\boldsymbol{\zeta})\boldsymbol{\varphi}\|_{\boldsymbol{\mathcal{Y}}}\|\boldsymbol{\xi} - \boldsymbol{\zeta}\|_{\boldsymbol{\Lambda}(\boldsymbol{\xi})}, \tag{7.86}$$

where the constant is given by $C = \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}\boldsymbol{L}_{\boldsymbol{\Lambda}}^{\boldsymbol{\mathcal{Y}}}$.

Proof. Let $\boldsymbol{\xi}, \boldsymbol{\zeta} \in B_{\mathcal{Y}}(\boldsymbol{R})$ and $\boldsymbol{\varphi}, \boldsymbol{\psi} \in \boldsymbol{\mathcal{X}}$. We use the Lipschitz continuity (4.4b) of $\boldsymbol{\Lambda}$ and the definition (7.72) of $\boldsymbol{\mathfrak{D}}$ to get

$$((\boldsymbol{\mathfrak{D}}(\boldsymbol{\xi}) - \boldsymbol{\mathfrak{D}}(\boldsymbol{\zeta}))\boldsymbol{\varphi} \mid \boldsymbol{\psi})_{\boldsymbol{\Lambda}(\boldsymbol{\xi})} = ((\boldsymbol{\Lambda}(\boldsymbol{\zeta}) - \boldsymbol{\Lambda}(\boldsymbol{\xi}))\boldsymbol{\mathfrak{D}}(\boldsymbol{\zeta})\boldsymbol{\varphi} \mid \boldsymbol{\psi})_{\boldsymbol{\mathcal{X}}} \leq \boldsymbol{L}_{\boldsymbol{\Lambda}}^{\mathcal{Y}} \|\boldsymbol{\zeta} - \boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}} \|\boldsymbol{\mathfrak{D}}(\boldsymbol{\zeta})\boldsymbol{\varphi}\|_{\boldsymbol{\mathcal{Y}}} \|\boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}}. \tag{7.87}$$

Hence, taking the supremum over all $\boldsymbol{\psi} \in \boldsymbol{\mathcal{X}}$ with $\|\boldsymbol{\psi}\|_{\boldsymbol{\Lambda}(\boldsymbol{\xi})} = 1$ together with (4.10) yields (7.86). □

In the following lemma, we estimate the defect $\widehat{\boldsymbol{d}}_{n+1}$. To do so, we first define the projection $\mathrm{P}_{\mathcal{H}} \in \mathcal{L}(\mathcal{X})$ by

$$\mathrm{P}_{\mathcal{H}}\, \psi := \begin{pmatrix} 0 \\ \psi_{\mathcal{H}} \end{pmatrix}, \qquad \psi = \begin{pmatrix} \psi_{\mathcal{V}} \\ \psi_{\mathcal{H}} \end{pmatrix} \in \mathcal{X}. \tag{7.88}$$

Note that the proof is based on [Hochbruck and Sturm, 2016, Lem. 5.2], where the defects arising in the error analysis for the full discretization of the linear Maxwell equations with a locally implicit method are analyzed.

Lemma 7.23. *Let $\boldsymbol{R}^{\partial_t} > 0$ such that*

$$\boldsymbol{z} \in C^1(J_T, B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}^{\partial_t})) \cap C(J_T, B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}))$$

holds. Furthermore, for $\tau > 0$, $K \in \mathbb{N}$, and $s_1, \ldots, s_K \in J_T$, let $0 \leq s_1 < \ldots < s_K \leq T$ and $|s_{k+1} - s_k| \leq \tau$, for $k = 1, \ldots, K-1$. If (7.69) is satisfied, we have for $z : J_T \to \mathcal{X}$, with $\mathrm{P}_{\mathcal{H}}\, z \in C^2(J_T, \mathcal{Y}_{\mathcal{V}})$ the bound

$$\begin{aligned} &\Big\| \sum_{j=0}^{K-1} \Big(\prod_{k=j+1}^{K-1} \widehat{\boldsymbol{\mathfrak{R}}}(\boldsymbol{z}_k) \Big) \widehat{\boldsymbol{\mathfrak{R}}}_-(\boldsymbol{z}_j)^{-1} \boldsymbol{\mathfrak{D}}(\boldsymbol{z}_j) \boldsymbol{\mathcal{J}}(z_{j+1} - z_j) \Big\|_{\boldsymbol{\Lambda}(\boldsymbol{z}_K)} \\ &\quad \leq C C_{\mathrm{stb}}^{\frac{1}{2}} (1 + K\tau) \tau \mathrm{e}^{C' s_K} \Big(\sup_{[s_0, s_K]} \|\mathcal{R}_{\mathrm{A}}\, \mathrm{P}_{\mathcal{H}}\, \partial_t^2 z\|_{\boldsymbol{\mathcal{X}}} + \sup_{[s_0, s_K]} \|\mathrm{A}\, \mathrm{P}_{\mathcal{H}}\, \partial_t^2 z\|_{\mathcal{X}} \Big), \end{aligned} \tag{7.89}$$

where we used the notation $\boldsymbol{z}_k = \boldsymbol{z}(s_k)$ and $z_k = z(s_k)$, for $k = 1, \ldots, K$.

Proof. Let $\boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$ and $\xi \in \mathcal{Y}$. From the definitions (7.72) and (7.73) of $\boldsymbol{\mathfrak{D}}(\boldsymbol{\xi})$, $\widehat{\boldsymbol{\mathfrak{R}}}_+(\boldsymbol{\xi})$, and $\widehat{\boldsymbol{\mathfrak{R}}}_-(\boldsymbol{\xi})$, we obtain

$$\boldsymbol{\mathfrak{D}}(\boldsymbol{\xi}) \boldsymbol{\mathcal{J}} \xi = \tau \boldsymbol{\mathcal{A}}(\boldsymbol{\xi}) \mathbf{A} \boldsymbol{\mathcal{J}}\, \mathrm{P}_{\mathcal{H}}\, \xi = (\widehat{\boldsymbol{\mathfrak{R}}}_+(\boldsymbol{\xi}) - \widehat{\boldsymbol{\mathfrak{R}}}_-(\boldsymbol{\xi})) \mathbf{A} \boldsymbol{\mathcal{J}}\, \mathrm{P}_{\mathcal{H}}\, \xi.$$

Thus, we further obtain

$$\widehat{\boldsymbol{\mathfrak{R}}}_-(\boldsymbol{\xi})^{-1} \boldsymbol{\mathfrak{D}}(\boldsymbol{\xi}) \boldsymbol{\mathcal{J}} \xi = (\widehat{\boldsymbol{\mathfrak{R}}}(\boldsymbol{\xi}) - \mathrm{Id}) \mathbf{A} \boldsymbol{\mathcal{J}}\, \mathrm{P}_{\mathcal{H}}\, \xi. \tag{7.90}$$

Hence, using (7.90) with $\boldsymbol{\xi} = \boldsymbol{z}_j$ and $\xi = z_{j+1} - z_j$ yields for $\triangle z_{j+1/2} := z_{j+1} - z_j$

$$\widehat{\boldsymbol{\mathfrak{R}}}_-(\boldsymbol{z}_j)^{-1} \boldsymbol{\mathfrak{D}}(\boldsymbol{z}_j) \boldsymbol{\mathcal{J}}(z_{j+1} - z_j) = (\widehat{\boldsymbol{\mathfrak{R}}}(\boldsymbol{z}_j) - \mathrm{Id}) \mathbf{A} \boldsymbol{\mathcal{J}}\, \mathrm{P}_{\mathcal{H}}\, \triangle z_{j+1/2}, \qquad j = 0, \ldots, K-1.$$

This yields

$$
\begin{aligned}
2\sum_{j=0}^{K-1}&\Big(\prod_{k=j+1}^{K-1}\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\Big)\widehat{\mathfrak{R}}_-(\boldsymbol{z}_j)^{-1}\mathfrak{D}(\boldsymbol{z}_j)\boldsymbol{\mathcal{J}}(z_{j+1}-z_j)\\
&=\sum_{j=0}^{K-1}\Big(\prod_{k=j+1}^{K-1}\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\Big)(\widehat{\mathfrak{R}}(\boldsymbol{z}_j)-\mathrm{Id})\mathbf{A}\boldsymbol{\mathcal{J}}\,\mathrm{P}_{\mathcal{H}}\,\triangle z_{j+1/2}\\
&=\sum_{j=0}^{K-1}\Big(\prod_{k=j}^{K-1}\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\Big)\mathbf{A}\boldsymbol{\mathcal{J}}\,\mathrm{P}_{\mathcal{H}}\,\triangle z_{j+1/2}-\sum_{j=0}^{K-1}\Big(\prod_{k=j+1}^{K-1}\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\Big)\mathbf{A}\boldsymbol{\mathcal{J}}\,\mathrm{P}_{\mathcal{H}}\,\triangle z_{j+1/2}\\
&=\sum_{j=-1}^{K-2}\Big(\prod_{k=j+1}^{K-1}\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\Big)\mathbf{A}\boldsymbol{\mathcal{J}}\,\mathrm{P}_{\mathcal{H}}\,\triangle z_{j+3/2}-\sum_{j=0}^{K-1}\Big(\prod_{k=j+1}^{K-1}\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\Big)\mathbf{A}\boldsymbol{\mathcal{J}}\,\mathrm{P}_{\mathcal{H}}\,\triangle z_{j+1/2}\\
&=\sum_{j=0}^{K-2}\Big(\prod_{k=j+1}^{K-1}\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\Big)\mathbf{A}\boldsymbol{\mathcal{J}}\,\mathrm{P}_{\mathcal{H}}(\triangle z_{j+3/2}-\triangle z_{j+1/2})\\
&\quad+\Big(\prod_{k=0}^{K-1}\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\Big)\mathbf{A}\boldsymbol{\mathcal{J}}\,\mathrm{P}_{\mathcal{H}}\,\triangle z_{1/2}-\mathbf{A}\boldsymbol{\mathcal{J}}\,\mathrm{P}_{\mathcal{H}}\,\triangle z_{K-1/2}.
\end{aligned}
\tag{7.91}
$$

We bound the norm of these three terms separately. First of all, we have

$$\triangle z_{j+1/2}=\int_{s_j}^{s_{j+1}}\partial_t z(s)\,\mathrm{d}s.$$

Thus, the norm equivalence (4.10) implies for the last term of (7.91)

$$\|\mathbf{A}\boldsymbol{\mathcal{J}}\,\mathrm{P}_{\mathcal{H}}\,\triangle z_{K-1/2}\|_{\mathbf{\Lambda}(\boldsymbol{z}_K)}\le C\tau\sup_{[s_{K-1},s_K]}\|\mathbf{A}\boldsymbol{\mathcal{J}}\,\mathrm{P}_{\mathcal{H}}\,\partial_t z\|_{\boldsymbol{\mathcal{X}}}.$$

For the second term of (7.91), we use (4.11) and Lemma 7.21 together with the arguments for the previous step to get

$$\Big\|\Big(\prod_{k=0}^{K-1}\widehat{\mathfrak{R}}(\boldsymbol{z}_k)\Big)\mathbf{A}\boldsymbol{\mathcal{J}}\,\mathrm{P}_{\mathcal{H}}\,\triangle z_{1/2}\Big\|_{\mathbf{\Lambda}(\boldsymbol{z}_K)}\le CC_{\mathrm{stb}}^{\frac{1}{2}}\tau\mathrm{e}^{C's_K}\sup_{[s_0,s_1]}\|\mathbf{A}\boldsymbol{\mathcal{J}}\,\mathrm{P}_{\mathcal{H}}\,\partial_t z\|_{\boldsymbol{\mathcal{X}}}. \tag{7.92}$$

Furthermore, Taylor's theorem yields

$$
\begin{aligned}
\triangle z_{j+3/2}&=\tau\partial_t z(s_{j+2})-\int_{s_{j+1}}^{s_{j+2}}|s_{j+1}-t|\,\partial_t^2 z(t)\,\mathrm{d}t,\\
\triangle z_{j+1/2}&=-\tau\partial_t z(s_j)-\int_{s_j}^{s_{j+1}}|s_{j+1}-t|\,\partial_t^2 z(t)\,\mathrm{d}t.
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
\triangle z_{j+3/2}-\triangle z_{j+1/2}&=\tau\big(\partial_t z(s_{j+2})-\partial_t z(s_j)-\int_{s_j}^{s_{j+2}}|s_{j+1}-t|\,\partial_t^2 z(t)\,\mathrm{d}t\\
&=\tau\int_{s_j}^{s_{j+2}}\Big(1-\frac{|s_{j+1}-t|}{\tau}\Big)\,\partial_t^2 z(t)\,\mathrm{d}t,
\end{aligned}
$$

which further implies

$$\|\mathbf{A}\boldsymbol{\mathcal{J}}\,\mathrm{P}_{\mathcal{H}}(\triangle z_{j+3/2}-\triangle z_{j+1/2})\|_{\boldsymbol{\mathcal{X}}}\le C\tau^2\big(\sup_{[s_j,s_{j+2}]}\|\mathbf{A}\boldsymbol{\mathcal{J}}\,\mathrm{P}_{\mathcal{H}}\,\partial_t^2 z\|_{\boldsymbol{\mathcal{X}}}\big).$$

Finally, for the first term of (7.91) we obtain as in (7.92)

$$\Big\| \sum_{j=0}^{K-2} \Big(\prod_{k=j+1}^{K-1} \widehat{\mathfrak{R}}(\boldsymbol{z}_K) \Big) \mathbf{A} \boldsymbol{\mathcal{J}} \, \mathrm{P}_{\mathcal{H}} (\triangle z_{j+3/2} - \triangle z_{j+1/2}) \Big\|_{\boldsymbol{z}_K} \leq C C_{\mathrm{stb}}^{\frac{1}{2}} K \tau^2 \mathrm{e}^{C' s_K} \sup_{[s_0, s_K]} \| \mathbf{A} \boldsymbol{\mathcal{J}} \, \mathrm{P}_{\mathcal{H}} \, \partial_t^2 z \|_{\boldsymbol{\mathcal{X}}}.$$

Since the definition (4.34) of the remainder $\mathcal{R}_{\mathrm{A}}$ and the boundedness (4.22) of $\boldsymbol{\mathcal{L}}_{\mathcal{X}}^*$ imply

$$\| \mathbf{A} \boldsymbol{\mathcal{J}} \varphi \|_{\boldsymbol{\mathcal{X}}} \leq \| \mathcal{R}_{\mathrm{A}} \varphi \|_{\boldsymbol{\mathcal{X}}} + C \| \mathrm{A} \varphi \|_{\mathcal{X}},$$

this concludes the proof. □

In the following theorem, we finally prove the wellposedness of the leapfrog scheme as well as an error estimate. Note that the following result is based on both the classical CFL condition (7.75) of the leapfrog scheme as well as the step size restriction (7.14) inherited from the problem, i.e., we assume the existence of constants $\varepsilon_0, C_0 > 0$ such that

$$\tau \leq \min\{ \boldsymbol{C}_{\mathbf{\Lambda}}(h)^{-1} \vartheta \boldsymbol{c}_{\mathbf{\Lambda}}, C_{\max}(h)^{-\frac{1}{2}} C_0 h^{\varepsilon_0} \} \tag{7.93}$$

holds. However, we emphasize that these restrictions do not accumulate, but act independently of each other.

Theorem 7.24. *Let Assumption 4.22, Assumption 7.2, and (7.69) be true. Furthermore, let $T < t^*(y_0)$ and $\vartheta \in (0,1)$ arbitrary but fixed. Then, there exist $h_0, \tau_0 > 0$ such that for all $h < h_0$ and $\tau < \tau_0$ under the combined step size restriction (7.93), the leapfrog scheme (7.70) is wellposed and satisfies for $n = 0, \ldots, N$ the estimate*

$$\begin{aligned} \| \boldsymbol{e}_n \|_{\mathbf{\Lambda}(\boldsymbol{\mathcal{I}} \widetilde{y}_n)} \leq C C_{\mathrm{stb}}^{\frac{3}{2}} (1 + t_n) \mathrm{e}^{C C_{\mathrm{stb}} t_n} \Big(& \| \boldsymbol{e}_0 \|_{\mathbf{\Lambda}(\boldsymbol{\mathcal{I}} \widetilde{y}_0)} + \sup_{[0,t_n]} \| (\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}}) y \|_{\boldsymbol{\mathcal{X}}} + \tau^2 C_{\tau^2} \\ & + \sup_{s_1, s_2 \in [0,t_n]} \| \mathcal{R}_\Lambda (y(s_1)) \partial_t y(s_2) \|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_n]} \| \mathcal{R}_{\mathrm{A}} y \|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_n]} \| \mathcal{R}_{\mathrm{F}}(\cdot, y) \|_{\boldsymbol{\mathcal{X}}} \Big), \end{aligned} \tag{7.94}$$

with a constant $C > 0$ independent of τ and h, and

$$C_{\tau^2} = \sup_{[0,t_n]} \| \partial_t y \|_{\mathcal{Y}} + \sup_{[0,t_n]} \| \partial_t^2 y \|_{\mathcal{Y}} + \sup_{[0,t_n]} \| \partial_t^3 y \|_{\mathcal{X}} + \sup_{[0,t_n]} \| \mathcal{R}_{\mathrm{A}} \, \mathrm{P}_{\mathcal{H}} \, \partial_t^2 y \|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_n]} \| \mathrm{A} \, \mathrm{P}_{\mathcal{H}} \, \partial_t^2 y \|_{\mathcal{X}}. \tag{7.95}$$

Before proving this theorem, we first state the following preliminary lemma.

Lemma 7.25. *Let the assumptions of Theorem 7.24 be satisfied and $\eta \leq N$. Further, assume that the first η steps of the leapfrog scheme (7.70) are wellposed and (7.94) is true for $n = 0, \ldots, \eta$. Then, there exist $h_0, \tau_0 > 0$ and a constant $\boldsymbol{R}^{\mathfrak{D}} > 0$ independent of h, τ, and η such that for all $h < h_0$, $\tau < \tau_0$ under the combined step size restriction (7.93)*

$$\| \widehat{\mathfrak{R}}_{-} (\boldsymbol{\mathcal{I}} \widetilde{y}_{n+1/2})^{-1} \underline{\mathfrak{D}}_{n+1/2} (\boldsymbol{y}_{n+1} - \boldsymbol{y}_n) \|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}^{\mathfrak{D}}$$

holds for $n = 0, \ldots, \eta - 1$.

Proof. Let $n \in \{0, \dots, \eta - 1\}$. Then, the definition (7.79) of the discrete error implies

$$\begin{aligned}\|\widehat{\mathfrak{R}}_-(\mathcal{I}\widetilde{y}_{n+1/2})^{-1}\underline{\mathfrak{D}}_{n+1/2}(\boldsymbol{y}_{n+1} - \boldsymbol{y}_n)\|_{\boldsymbol{\mathcal{Y}}} \leq \|\widehat{\mathfrak{R}}_-(\mathcal{I}\widetilde{y}_{n+1/2})^{-1}\underline{\mathfrak{D}}_{n+1/2}(\boldsymbol{e}_{n+1} + \boldsymbol{e}_n)\|_{\boldsymbol{\mathcal{Y}}} \\ + \|\widehat{\mathfrak{R}}_-(\mathcal{I}\widetilde{y}_{n+1/2})^{-1}\underline{\mathfrak{D}}_{n+1/2}\mathcal{J}(\widetilde{y}_{n+1} - \widetilde{y}_n)\|_{\boldsymbol{\mathcal{Y}}}.\end{aligned} \tag{7.96}$$

From the inverse estimate (4.1), the definition (7.72) of $\mathfrak{D}$, the norm equivalence (4.10), and the boundedness (7.76d) of $\widehat{\mathfrak{R}}_-$, we derive

$$\|\widehat{\mathfrak{R}}_-(\mathcal{I}\widetilde{y}_{n+1/2})^{-1}\underline{\mathfrak{D}}_{n+1/2}(\boldsymbol{e}_{n+1} + \boldsymbol{e}_n)\|_{\boldsymbol{\mathcal{Y}}} \leq CC_{\mathrm{stb}}\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\tfrac{\tau}{2}\|\underline{\boldsymbol{\mathcal{A}}}_{n+1/2}\mathbf{A}\,\boldsymbol{P}_{\mathcal{H}}(\boldsymbol{e}_{n+1} + \boldsymbol{e}_n)\|_{\boldsymbol{\mathcal{X}}},$$

where the projection $\boldsymbol{P}_{\mathcal{H}} \in \mathcal{L}(\boldsymbol{\mathcal{X}})$ is defined corresponding to (7.88). Thus, the boundedness (7.77) of $\mathbf{A}_{\boldsymbol{\mathcal{V}}}$ and $\mathbf{A}_{\boldsymbol{\mathcal{H}}}$ together with the CFL condition (7.75) imply for the first term of (7.96)

$$\|\widehat{\mathfrak{R}}_-(\mathcal{I}\widetilde{y}_{n+1/2})^{-1}\underline{\mathfrak{D}}_{n+1/2}(\boldsymbol{e}_{n+1} + \boldsymbol{e}_n)\|_{\boldsymbol{\mathcal{Y}}} \leq CC_{\mathrm{stb}}\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\vartheta \boldsymbol{C}_{\mathbf{A}}(h)(\|\boldsymbol{e}_{n+1}\|_{\boldsymbol{\mathcal{X}}} + \|\boldsymbol{e}_n\|_{\boldsymbol{\mathcal{X}}}). \tag{7.97}$$

Furthermore, (7.72), the boundedness (4.3) of $\boldsymbol{\Lambda}$, and the definition (4.9) of the state-dependent norm yield

$$\begin{aligned}&\|\widehat{\mathfrak{R}}_-(\mathcal{I}\widetilde{y}_{n+1/2})^{-1}\underline{\mathfrak{D}}_{n+1/2}\mathcal{J}(\widetilde{y}_{n+1} - \widetilde{y}_n)\|_{\boldsymbol{\mathcal{Y}}} \\ &\quad \leq C\boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\|\widehat{\mathfrak{R}}_-(\mathcal{I}\widetilde{y}_{n+1/2})^{-1}\widetilde{\mathfrak{D}}_{n+1/2}\mathcal{J}(\widetilde{y}_{n+1} - \widetilde{y}_n)\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{n+1/2})}.\end{aligned}$$

Hence, Lemma 7.23 implies

$$\begin{aligned}&\|\widehat{\mathfrak{R}}_-(\mathcal{I}\widetilde{y}_{n+1/2})^{-1}\underline{\mathfrak{D}}_{n+1/2}\mathcal{J}(\widetilde{y}_{n+1} - \widetilde{y}_n)\|_{\boldsymbol{\mathcal{Y}}} \\ &\quad \leq C\boldsymbol{C}_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}}(h)\tau\big(\sup_{[t_n,t_{n+1}]}\|\mathcal{R}_{\mathrm{A}}\,\mathrm{P}_{\mathcal{H}}\,\partial_t^2 y\|_{\boldsymbol{\mathcal{X}}} + \sup_{[t_n,t_{n+1}]}\|\mathrm{A}\,\mathrm{P}_{\mathcal{H}}\,\partial_t^2 y\|_{\mathcal{X}}\big).\end{aligned} \tag{7.98}$$

Finally, using (7.97) and (7.98) in (7.96), we get

$$\begin{aligned}\|\underline{\mathfrak{D}}_{n+1/2}(\boldsymbol{y}_{n+1} - \boldsymbol{y}_n)\|_{\boldsymbol{\mathcal{Y}}} \leq CC_{\mathrm{stb}}C_{\max}(h)\Big(&\vartheta(\|\boldsymbol{e}_{n+1}\|_{\boldsymbol{\mathcal{X}}} + \|\boldsymbol{e}_n\|_{\boldsymbol{\mathcal{X}}}) \\ &+ \tau\big(\sup_{[t_n,t_{n+1}]}\|\mathcal{R}_{\mathrm{A}}\,\mathrm{P}_{\mathcal{H}}\,\partial_t^2 y\|_{\boldsymbol{\mathcal{X}}} + \sup_{[t_n,t_{n+1}]}\|\mathrm{A}\,\mathrm{P}_{\mathcal{H}}\,\partial_t^2 y\|_{\mathcal{X}}\big)\Big).\end{aligned}$$

Thus, the error estimate (7.94), Assumption 4.22 and the step size restriction (7.14) yield

$$\|\underline{\mathfrak{D}}_{n+1/2}(\boldsymbol{y}_{n+1} - \boldsymbol{y}_n)\|_{\boldsymbol{\mathcal{Y}}} \to 0$$

uniformly in n, for $\tau, h \to 0$ under the combined step size restriction (7.93). This proves the result. □

Proof of Theorem 7.24. The proof essentially follows the same approach as the proofs of Theorem 7.3 for the full discretization with the implicit midpoint rules, i.e., we show the statement by induction, as we alternately prove the error bound (7.94) and the wellposedness of the next step.

For the induction base ($n = 0$), Assumption 4.22 yields the existence of $h_1 > 0$ such that

$$\|\boldsymbol{y}_0\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R} + \boldsymbol{C}_{\mathcal{I}}R), \qquad \|\boldsymbol{\mathcal{A}}(\boldsymbol{y}_0)\boldsymbol{y}_0\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \boldsymbol{C}_{\mathcal{I}}R^{\mathcal{A}})$$

holds for all $h < h_1$. Moreover, (7.94) is trivially satisfied in this case.

For the induction step, we assume for some $\eta \in \{0,\dots,N-1\}$ arbitrary but fixed, that the first $\eta \geq 0$ steps of the leapfrog scheme are well defined. More precisely, we assume for $n = 0,\dots,\eta$, that

$$\|\boldsymbol{y}_n\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}} R), \qquad \|\boldsymbol{\mathcal{A}}_n \boldsymbol{y}_n\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \boldsymbol{C}_{\boldsymbol{\mathcal{I}}} R^{\mathcal{A}}), \tag{7.99}$$

as well as (7.94) is satisfied. Then, Lemma 7.25 implies the existence of a constant $\boldsymbol{R}^{\boldsymbol{\mathfrak{D}}} > 0$ such that

$$\|\widehat{\boldsymbol{\mathfrak{R}}}_-(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})^{-1}\underline{\boldsymbol{\mathfrak{D}}}_{n+1/2}(\boldsymbol{y}_{n+1} - \boldsymbol{y}_n)\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}^{\boldsymbol{\mathfrak{D}}}$$

holds for $n = 0,\dots,\eta-1$. In the following, we prove that (7.99) is also true for $n = \eta+1$.

Concerning the wellposedness of the next step, the definition (7.70) of the leapfrog scheme directly yields the existence of $\boldsymbol{y}_{\eta+1} \in \boldsymbol{\mathcal{X}}$. Furthermore, (7.70) together with the boundedness (7.77) and (4.6) of $\mathbf{A}_{\mathcal{V}}$, $\mathbf{A}_{\mathcal{H}}$, and $\mathbf{F}_{\mathcal{V}}$, $\mathbf{F}_{\mathcal{H}}$, respectively, imply the existence of $\tau_{0,\mathrm{w}} > 0$ such that

$$\begin{gathered}\|\boldsymbol{y}_{\eta+1}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}, \qquad \|\underline{\boldsymbol{\mathcal{A}}}_{\eta+1/2}\boldsymbol{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}^{\boldsymbol{\mathcal{A}}}, \\ \|\widehat{\boldsymbol{\mathfrak{R}}}_-(\boldsymbol{\mathcal{I}}\widetilde{y}_{\eta+1/2})^{-1}\underline{\boldsymbol{\mathfrak{D}}}_{\eta+1/2}(\boldsymbol{y}_{\eta+1} - \boldsymbol{y}_\eta)\|_{\boldsymbol{\mathcal{Y}}} < \boldsymbol{R}^{\boldsymbol{\mathfrak{D}}}\end{gathered} \tag{7.100}$$

holds for all $\tau < \tau_{0,\mathrm{w}}$. Note that $\tau_{0,\mathrm{w}}$ depends on the upper bound $h_0 > 0$ for the space discretization parameters. At the end of this proof, we see that this dependency is given by the combined step size restriction (7.93). Thus, let in the following $\tau < \tau_{0,\mathrm{w}}$.

We now bound the terms on the right-hand side of (7.85). As we already have the bounds (7.67) and (7.68) for $\boldsymbol{g}_{n+1}^{\mathrm{LI}}$, we focus now on $\boldsymbol{d}_{n+1}$ and $\widehat{\boldsymbol{d}}_{n+1}$. Based on the definition (7.84a) of $\boldsymbol{d}_{n+1}$, we first obtain from (7.86) the estimate

$$\begin{aligned}&\|\widehat{\boldsymbol{\mathfrak{R}}}_-(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})^{-1}\boldsymbol{d}_{n+1}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})} \\ &\qquad \leq C\|\widehat{\boldsymbol{\mathfrak{R}}}_-(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})^{-1}\underline{\boldsymbol{\mathfrak{D}}}_{n+1/2}(\boldsymbol{y}_{n+1} - \boldsymbol{y}_j)\|_{\boldsymbol{\mathcal{Y}}}\|\underline{\boldsymbol{y}}_{n+1/2} - \boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})}.\end{aligned}$$

Thus, (7.100) and the norm equivalence (4.10) imply

$$\|\widehat{\boldsymbol{\mathfrak{R}}}_-(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})^{-1}\boldsymbol{d}_{n+1}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})} \leq C\Big(\|\underline{\boldsymbol{y}}_{n+1/2} - \boldsymbol{\mathcal{J}}\widetilde{y}_{n+1/2}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})} + \|(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{I}})\widetilde{y}_{n+1/2}\|_{\boldsymbol{\mathcal{X}}}\Big),$$

with a constant $C > 0$ depending on $\boldsymbol{R}^{\boldsymbol{\mathfrak{D}}}$. Thus, we obtain from (7.36) for the case $n = 0$

$$\|\widehat{\boldsymbol{\mathfrak{R}}}_-(\boldsymbol{\mathcal{I}}\widetilde{y}_{1/2})^{-1}\boldsymbol{d}_1\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{1/2})} \leq C\Big(\|\boldsymbol{e}_0\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_0)} + \tau \sup_{[t_0,t_{1/2}]}\|\partial_t y\|_{\mathcal{Y}} + \|(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{I}})\widetilde{y}_{1/2}\|_{\boldsymbol{\mathcal{X}}}\Big).$$

We further obtain from (7.39) and (4.11) for the case $n > 0$

$$\begin{aligned}\|\widehat{\boldsymbol{\mathfrak{R}}}_-(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})^{-1}\boldsymbol{d}_{n+1}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})} &\leq C\Big(\mathrm{e}^{C'(t_{n+1/2}-t_n)}\|\boldsymbol{e}_n\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})} \\ &\quad + \mathrm{e}^{C'(t_{n+1/2}-t_{n-1})}\|\boldsymbol{e}_{n-1}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{n+1/2})} + \tau^2 \sup_{[t_{n-1},t_{n+1/2}]}\|\partial_t^2 y\|_{\mathcal{Y}} + \|(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{I}})\widetilde{y}_{n+1/2}\|_{\boldsymbol{\mathcal{X}}}\Big).\end{aligned}$$

Thus, the combination of these bounds with the corresponding bounds for $\boldsymbol{g}_{j+1}^{\mathrm{LI}}$ in (7.67) and (7.68) for $n = 0$ and $n > 0$, respectively, implies

$$\begin{aligned}\|\boldsymbol{g}_1^{\mathrm{LI}}\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{1/2})} + \|\widehat{\boldsymbol{\mathfrak{R}}}_-(\boldsymbol{\mathcal{I}}\widetilde{y}_{1/2})^{-1}\boldsymbol{d}_1\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_{1/2})} &\leq C\Big(\mathrm{e}^{C't_{1/2}}\|\boldsymbol{e}_0\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\widetilde{y}_0)} + \tau \sup_{[t_0,t_{1/2}]}\|\partial_t y\|_{\mathcal{Y}} + \tau^2 C_{\tau^2} \\ &\quad + \|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\widetilde{y}_{1/2}\|_{\boldsymbol{\mathcal{X}}} + \sup_{[t_0,t_1]}\|\mathcal{R}_\Lambda(\widetilde{y}_{1/2})\partial_t y\|_{\boldsymbol{\mathcal{X}}} + \|\mathcal{R}_{\mathrm{A}}(\widetilde{y}_1 + \widetilde{y}_0)\|_{\boldsymbol{\mathcal{X}}} + \|\mathcal{R}_{\mathrm{F}}(t_{1/2},\widetilde{y}_{1/2})\|_{\boldsymbol{\mathcal{X}}}\Big)\end{aligned} \tag{7.101}$$

and

$$
\begin{aligned}
\|\boldsymbol{g}_{n+1}^{\mathrm{LI}}\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{n+1/2})} + \|\widehat{\mathfrak{R}}_{-}(\mathcal{I}\widetilde{y}_{n+1/2})^{-1}\boldsymbol{d}_{n+1}\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{n+1/2})} &\leq C\Big(\mathrm{e}^{C'(t_{n+1/2}-t_n)}\|\boldsymbol{e}_n\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_n)} \\
&+ \mathrm{e}^{C'(t_{n+1/2}-t_{n-1})}\|\boldsymbol{e}_{n-1}\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{n-1})} + \|(\mathcal{I}-\mathcal{J})\widetilde{y}_{n+1/2}\|_{\boldsymbol{\mathcal{X}}} + \tau^2 C_{\tau^2} \\
&+ \sup_{[t_n,t_{n+1}]}\|\mathcal{R}_{\Lambda}(\widetilde{y}_{n+1/2})\partial_t y\|_{\boldsymbol{\mathcal{X}}} + \|\mathcal{R}_{\mathrm{A}}(\widetilde{y}_{n+1}+\widetilde{y}_n)\|_{\boldsymbol{\mathcal{X}}} + \|\mathcal{R}_{\mathrm{F}}(t_{n+1/2},\widetilde{y}_{n+1/2})\|_{\boldsymbol{\mathcal{X}}}\Big).
\end{aligned}
\tag{7.102}
$$

Furthermore, we obtain from the definition (7.84b) of $\widehat{\boldsymbol{d}}_{n+1}$ together with (7.89) and (4.11)

$$
\Big\|\sum_{j=0}^{n-1}\Big(\prod_{k=n+1}^{n-1}\widehat{\mathfrak{R}}(\mathcal{I}\widetilde{y}_{k+1/2})\Big)\widehat{\mathfrak{R}}_{-}(\mathcal{I}\widetilde{y}_{n+1/2})^{-1}\widehat{\boldsymbol{d}}_{n+1}\Big\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_n)} \leq C C_{\mathrm{stb}}^{\frac{1}{2}}(1+t_n)\mathrm{e}^{C' t_n}\tau C_{\tau}^{2}. \tag{7.103}
$$

Collecting these results, we use (7.101), (7.102), and (7.103) in (7.85) to obtain

$$
\begin{aligned}
\mathrm{e}^{-C' t_{\eta+1}}\|\boldsymbol{e}_{\eta+1}\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1})} &\leq C C_{\mathrm{stb}}^{\frac{1}{2}}\|\boldsymbol{e}_0\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_0)} + C C_{\mathrm{stb}}^{\frac{3}{2}}\tau\sum_{j=0}^{\eta}\mathrm{e}^{-C' t_j}\|\boldsymbol{e}_j\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_j)} \\
&+ C C_{\mathrm{stb}}^{\frac{3}{2}}(1+t_{\eta+1})\Big(\tau^2 C_{\tau^2} + \sup_{[0,t_{\eta+1}]}\|(\mathcal{I}-\mathcal{J})y\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_{\eta+1}]}\|\mathcal{R}_{\mathrm{A}} y\|_{\boldsymbol{\mathcal{X}}} \\
&+ \sup_{s_1,s_2\in[0,t_{\eta+1}]}\|\mathcal{R}_{\Lambda}(y(s_1))\partial_t y(s_2)\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_{\eta+1}]}\|\mathcal{R}_{\mathrm{F}}(\cdot,y)\|_{\boldsymbol{\mathcal{X}}}\Big).
\end{aligned}
$$

Application of the discrete Gronwall inequality finally yields (7.94).

To conclude, we have to prove that (7.99) holds with $n=\eta+1$ for $\tau_{0,\mathrm{w}}$ given by the combined step size restriction (7.93). We also have to show that this step size restriction is sufficient to ensure (7.100).

Note that due to (7.94) and Assumption 4.22, Lemma 4.24 yields the existence of $\tau_0, h_0 > 0$ with $h_0 \leq h_1$ such that (7.99) as well as the first two bounds in (7.100) are satisfied for all $\tau < \tau_0$ and $h < h_0$ under the combined step size restriction (7.93). As the last bound follows from Lemma 7.25 under the same condition, this concludes the proof. □

Analogous to Corollary 7.4 for the implicit midpoint rules, Theorem 7.24 yields the following convergence result.

Corollary 7.26. *Under the assumptions of Theorem 7.24, the error of the leapfrog scheme (7.70) is for $n = 0, \dots, N$ bounded by*

$$
\begin{aligned}
\|y(t_n) - \boldsymbol{\mathcal{L}}\boldsymbol{y}_n\|_{\mathcal{X}} &\leq \|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{J}})y(t_n)\|_{\mathcal{X}} + C C_{\mathrm{stb}}^{\frac{3}{2}}(1+t_n)\mathrm{e}^{C C_{\mathrm{stb}} t_n}\Big(\|\boldsymbol{\mathcal{J}} y_0 - \boldsymbol{y}_0\|_{\boldsymbol{\mathcal{X}}} + \tau^2 C_{\tau^2} \\
&+ \sup_{[0,t_n]}\|(\mathcal{I}-\mathcal{J})y\|_{\boldsymbol{\mathcal{X}}} + \sup_{s_1,s_2\in[0,t_n]}\|\mathcal{R}_{\Lambda}(y(s_1))\partial_t y(s_2)\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_n]}\|\mathcal{R}_{\mathrm{A}} y\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_n]}\|\mathcal{R}_{\mathrm{F}}(\cdot,y)\|_{\boldsymbol{\mathcal{X}}}\Big),
\end{aligned}
$$

with a constant $C > 0$ independent of τ and h and C_{τ^2} from (7.95). Furthermore, we have

$$
\|y(t_n) - \boldsymbol{\mathcal{L}}\boldsymbol{y}_n\|_{\mathcal{X}} \to 0, \qquad n = 0, \dots, N,
$$

for $\tau, h \to 0$ satisfying the step size restriction (7.14).

Based on the discretization of local nonlinearities discussed in Section 4.4, we further directly obtain the following error bound analogous to Corollary 7.5.

Corollary 7.27. *Let Assumption 4.26 be satisfied. Then, the statements of Theorem 7.24 and Corollary 7.26 are also valid if we replace Assumption 4.22 by Assumption 4.28. In particular, the error of the leapfrog scheme (7.70) is for $n = 0, \dots, N$ bounded by*

$$\begin{aligned}\|y(t_n) - \boldsymbol{\mathcal{L}} \boldsymbol{y}_n\|_{\mathcal{X}} \leq \|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{J}})y(t_n)\|_{\mathcal{X}} + C C_{\mathrm{stb}}^{\frac{3}{2}}(1+t_n)\mathrm{e}^{CC_{\mathrm{stb}}t_n}\Big(&\|\boldsymbol{\mathcal{J}}y_0 - \boldsymbol{y}_0\|_{\boldsymbol{\mathcal{X}}} + \sup_{[0,t_n]}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})y\|_{\boldsymbol{\mathcal{X}}} \\ &+ \tau^2 C_{\tau^2} + \sup_{[0,t_n]}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\partial_t y\|_{\boldsymbol{\mathcal{X}}} + \sup_{s_1,s_2\in[0,t_n]}\|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}})\Lambda(y(s_1))\partial_t y(s_2)\|_{\mathcal{X}} \\ &+ \sup_{s_1,s_2\in[0,t_n]} \triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}\Lambda(y(s_1))\partial_t y(s_2)) + \sup_{[0,t_n]}\|\mathcal{R}_{\mathrm{A}} y\|_{\boldsymbol{\mathcal{X}}} \\ &+ \sup_{[0,t_n]}\|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}})\mathrm{F}(\cdot, y)\|_{\mathcal{X}} + \sup_{[0,t_n]} \triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}\mathrm{F}(\cdot, y))\Big),\end{aligned} \tag{7.104}$$

with a constant $C > 0$ independent of τ and h and C_{τ^2} from (7.95).

We conclude this subsection with the following remark on the necessity of the additional assumption (7.69).

Remark 7.28. *Note that (7.69) is not necessary to prove Lemma 7.20 and Lemma 7.23, as equivalent results also hold without this assumption. However, the assumption (7.69) is essential for Lemma 7.21 and Lemma 7.22, as (7.78) and (7.87) depend on $\boldsymbol{\Lambda}(\boldsymbol{\xi})\boldsymbol{\mathfrak{D}}(\xi) = \frac{\tau}{2}\boldsymbol{P}_{\mathcal{H}}\mathbf{A}^2$ being independent of $\boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$.*

CHAPTER 8

Full discretization of the specific examples

As for the space discretization, we now apply the abstract results for the full discretization of quasilinear wave-type equations presented in Chapter 7 to the specific examples from Chapter 3.

8.1 Example: Westervelt equation

Based on the space discretization of the Westervelt equation introduced in Section 5.1, we now investigate the full discretization with the implicit midpoint rules and the leapfrog scheme. In particular, we combine the arguments from the error analysis for the space discretization of the Westervelt equation in Theorem 5.3 with the abstract error estimates for the implicit midpoint rules in Corollary 7.5 and the leapfrog scheme in Corollary 7.27.

8.1.1 Example: Westervelt equation (1D)

In the following, we continue the analysis of the one-dimensional Westervelt equation. In particular, based on the space discretization presented in Section 5.1.1, we now provide error estimates for the full discretization.

Theorem 8.1. *Let $d = 1$ and $p \geq 2$. For $\boldsymbol{R} \in (0, \frac{1}{|\varkappa|})$, let the assumptions of* Theorem 3.3 *be satisfied with $R < \boldsymbol{C_{\mathcal{I}} R}$. Further, let the solution u of* (3.7) *satisfy*

$$u \in C^4(J_T, L^2(\Omega)) \cap C^3(J_T, H^1(\Omega)) \cap C^2(J_T, H^p(\Omega)) \cap C^1(J_T, H^{p+1}(\Omega)) \cap C(J_T, L^\infty(R)).$$

Then, we obtain the following results for the full discretization of the one-dimensional Westervelt equation.

(i) If there are constants $\varepsilon_0, C_0 > 0$ such that the discretization parameters $h, \tau > 0$ satisfy

$$\tau \leq C_0 h^{\frac{3}{4}+\varepsilon_0}, \tag{8.1}$$

there exist $h_0, \tau_0 > 0$ such that for $h < h_0$ and $\tau < \tau_0$, the full discretization of (3.7) *with either of the implicit midpoint rules is well defined.*

(ii) Let $\vartheta \in (0,1)$ and

$$C_1 = 2C_{\mathrm{inv}}^{-1} \min\Big\{1, (1-|\varkappa|\boldsymbol{R})\frac{C_{norm}}{c_{norm}}\Big\}^{\frac{1}{2}}. \tag{8.2}$$

If the discretization parameters $h, \tau > 0$ satisfy

$$\tau \leq C_1 \vartheta h, \tag{8.3}$$

there exist $h_0, \tau_0 > 0$ such that for $h < h_0$ and $\tau < \tau_0$, the full discretization of (3.7) with the leapfrog scheme is well defined.

In both cases, the approximations $\boldsymbol{y}_n = (\boldsymbol{u}_n, \boldsymbol{v}_n) \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$ obtained by the application of either of the schemes satisfy for $n = 0, \dots, N$

$$|u(t_n) - \boldsymbol{u}_n|_{H^1(\Omega)} + \|\partial_t u(t_n) - \boldsymbol{v}_n\|_{L^2(\Omega)} \leq C_u(1+t_n)\mathrm{e}^{Ct_n}(h^p + \tau^2), \tag{8.4}$$

with constants $C_u, C > 0$ independent of h, τ, and T, but C_u depending on the solution u and its derivatives.

Proof. As we have shown in the proof of Theorem 5.3, the discretization of the one-dimensional Westervelt equation satisfies Assumption 4.28. Furthermore, we obtain from the representation (5.23) of $C_{\max}(h)$, that the step size restriction (7.14) is satisfied if (8.1) holds.

Moreover, the representations (5.17) and (5.22) of $\boldsymbol{c}_{\boldsymbol{\Lambda}}$ and $\boldsymbol{C}_{\boldsymbol{\Lambda}}(h)$, respectively, yield that (7.75) corresponds to

$$\tau C_{\mathrm{inv}} h^{-1} \leq 2\vartheta \min\Big\{1, (1-|\varkappa|\boldsymbol{R})\frac{C_{\mathrm{norm}}}{c_{\mathrm{norm}}}\Big\}^{\frac{1}{2}}.$$

Hence, due to (5.23) and (8.2), (8.3) implies (7.93).

In the following we employ the abstract error estimates for the full discretization with $y = (u, \partial_t u)$ and $\boldsymbol{y} = (\boldsymbol{u}, \boldsymbol{v})$, starting with the implicit midpoint rules. As we consider a conforming discretization, Corollary 7.5 implies for either of the implicit midpoint rules the estimate

$$\begin{aligned}\|y(t_n) - \boldsymbol{y}_n\|_{\mathcal{X}} \leq \|(\mathrm{Id} - \boldsymbol{\mathcal{J}})y(t_n)\|_{\mathcal{X}} + C(1+t_n)\mathrm{e}^{Ct_n}\Big(\|\boldsymbol{\mathcal{J}}y_0 - \boldsymbol{y}_0\|_{\mathcal{X}} + \sup_{[0,t_n]}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})y\|_{\mathcal{X}} \\ + \tau^2\big(\sup_{[0,t_n]}\|\partial_t y\|_{\mathcal{Y}} + \sup_{[0,t_n]}\|\partial_t^2 y\|_{\mathcal{Y}} + \sup_{[0,t_n]}\|\partial_t^3 y\|_{\mathcal{X}}\big) + \sup_{[0,t_n]}\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\partial_t y\|_{\mathcal{X}} \\ + \sup_{s_1,s_2\in[0,t_n]}\|(\mathrm{Id} - \boldsymbol{\mathcal{I}})\Lambda(y(s_1))\partial_t y(s_2)\|_{\mathcal{X}} + \sup_{[0,t_n]}\|\mathcal{R}_{\mathrm{A}} y\|_{\mathcal{X}} + \sup_{[0,t_n]}\|(\mathrm{Id} - \boldsymbol{\mathcal{I}})\mathrm{F}(\cdot, y)\|_{\mathcal{X}}\Big).\end{aligned}$$

Thus, the definition (5.12) of the initial values, the estimate (5.24) for the difference between $\boldsymbol{\mathcal{I}}$ and $\boldsymbol{\mathcal{J}}$, and the bound (5.26) for the remainder $\mathcal{R}_{\mathrm{A}}$ yield

$$\begin{aligned}\|y(t_n) - \boldsymbol{y}_n\|_{\mathcal{X}} \leq C(1+t_n)\mathrm{e}^{Ct_n}\Big(\sup_{[0,t_n]}\|(\boldsymbol{\mathcal{I}} - \mathrm{Id})y\|_{\mathcal{X}} + \tau^2\big(\sup_{[0,t_n]}\|\partial_t y\|_{\mathcal{Y}} + \sup_{[0,t_n]}\|\partial_t^2 y\|_{\mathcal{Y}} + \sup_{[0,t_n]}\|\partial_t^3 y\|_{\mathcal{X}}\big) \\ + \sup_{[0,t_n]}\|(\boldsymbol{\mathcal{I}} - \mathrm{Id})\partial_t y\|_{\mathcal{X}} + \sup_{s_1,s_2\in[0,t_n]}\|(\mathrm{Id} - \boldsymbol{\mathcal{I}})\Lambda(y(s_1))\partial_t y(s_2)\|_{\mathcal{X}} + \sup_{[0,t_n]}\|(\mathrm{Id} - \boldsymbol{\mathcal{I}})\mathrm{F}(\cdot, y)\|_{\mathcal{X}}\Big).\end{aligned}$$

Hence, (5.19) implies

$$|u(t_n)-\boldsymbol{u}_n|_{H^1(\Omega)}+\|\partial_t u(t_n)-\boldsymbol{v}_n\|_{L^2(\Omega)} \leq C(1+t_n)\mathrm{e}^{Ct}\Big(h^p\Big(\sup_{[0,t_n]}\|u\|_{H^{p+1}(\Omega)}+\sup_{[0,t_n]}\|\partial_t u\|_{H^{p+1}(\Omega)}$$
$$+\sup_{[0,t_n]}\|\partial_t^2 u\|_{H^p(\Omega)}+\sup_{s_1,s_2\in[0,t_n]}\|\Lambda(y(s_1))\partial_t y(s_2)\|_{H^{p+1}(\Omega)\times H^p(\Omega)}+\sup_{[0,t_n]}\|\mathrm{F}(y)\|_{H^{p+1}(\Omega)\times H^p(\Omega)}\Big)$$
$$+\tau^2\Big(\sup_{[0,t_n]}\|\partial_t u\|_{H^2(\Omega)}+\sup_{[0,t_n]}\|\partial_t^2 u\|_{H^2(\Omega)}+\sup_{[0,t_n]}\|\partial_t^3 u\|_{H^1(\Omega)}+\sup_{[0,t_n]}\|\partial_t^4 u\|_{L^2(\Omega)}\Big)\Big).$$

Finally, (5.28) proves (8.4) for the implicit midpoint rules.

For the leapfrog scheme, we employ Corollary 7.27 . However, as most of the terms in (7.104) coincide with the terms for the implicit midpoint rules, we only focus on the derivation of a bound for the additional terms in C_{τ^2}, which is defined in (7.95). First, we obtain from (7.88), (5.25), and (5.20)

$$\sup_{[0,t_n]}\|\mathcal{R}_{\mathrm{A}}\,\mathrm{P}_{\mathcal{H}}\,\partial_t^2 y\|_{\boldsymbol{\mathcal{X}}}=\sup_{[0,t_n]}|(\boldsymbol{\mathcal{I}}_{\mathcal{V}}-\mathrm{Id})\partial_t^3 u|_{H^1(\Omega)}\leq C\sup_{[0,t_n]}|\partial_t^3 u|_{H^1(\Omega)}.$$

Second, (3.8) and (7.88) yield

$$\sup_{[0,t_n]}\|\mathrm{A}\,\mathrm{P}_{\mathcal{H}}\,\partial_t^2 y\|_{\mathcal{X}}=\sup_{[0,t_n]}|\partial_t^3 u|_{H^1(\Omega)}.$$

Thus, (7.104) together with the estimate (5.28) for the nonlinearities proves (8.4) for the leapfrog scheme. □

In Section 9.1 we present numerical experiments confirming these results.

8.1.2 Example: Westervelt equation (2D, 3D)

As in the previous section for the one-dimensional case, we now focus on the full discretization of the multi-dimensional Westervelt equation. In particular, we employ the space discretization from Section 5.1.2. Numerical results are shown in Section 9.2.

Theorem 8.2. *For $d\in\{2,3\}$ and $p\geq 3$, let $\Omega\subset\mathbb{R}^d$ be a bounded domain with C^{p+1}-boundary. For $\boldsymbol{R}\in(0,\frac{1}{|\varkappa|})$, let the assumptions of Theorem 3.3 be satisfied with $R<\boldsymbol{C}_{\boldsymbol{\mathcal{I}}}\boldsymbol{R}$. Further, let the solution u of (3.7) satisfy*

$$u\in C^4(J_T,L^2(\Omega))\cap C^3(J_T,H^2(\Omega))\cap C^2(J_T,H^p(\Omega))\cap C^1(J_T,H^{p+1}(\Omega))\cap C(J_T,L^\infty(R)).$$

Then, we obtain the following results for the full discretization of the multi-dimensional Westervelt equation.

(i) If there exist $\varepsilon_0, C_0>0$ such that the discretization parameters $h,\tau>0$ satisfy

$$\tau<C_0h^{\frac{d+2}{4}+\varepsilon_0}, \tag{8.5}$$

the full discretization of (3.7) with either of the implicit midpoint rules is well defined.

(ii) Let $\vartheta \in (0,1)$ and $C_1 > 0$ be given as in (8.2). If there exist $\varepsilon_0, C_0 > 0$ such that the discretization parameters $h, \tau > 0$ satisfy

$$\tau < \min\Big\{C_1\vartheta h, C_0 h^{\frac{d+2}{4}+\varepsilon_0}\Big\},$$

the full discretization of (3.7) with the leapfrog scheme is well defined.

In both cases the approximations $\boldsymbol{y}_n = (\boldsymbol{u}_n, \boldsymbol{v}_n) \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$ obtained by the application of either of the schemes satisfy for $n = 0, \dots, N$

$$|u(t_n) - \boldsymbol{\mathcal{L}}_{\mathcal{V}}\boldsymbol{u}_n|_{H^1(\Omega)} + \|\partial_t u(t_n) - \boldsymbol{\mathcal{L}}_{\mathcal{H}}\boldsymbol{v}_n\|_{L^2(\Omega)} \le C_u(1+t_n)\mathrm{e}^{Ct_n}(h^p + \tau^2), \tag{8.6}$$

with constants $C_u, C > 0$ independent of h, τ, and T, but C_u depending on the solution u and its derivatives.

Proof. As in the one-dimensional case, the statement follows by combining the space discretization in Section 5.1.2 with the abstract error estimates Corollary 7.5 and Corollary 7.27 for the full discretization with the implicit midpoint rules and the leapfrog scheme, respectively.

However, note that the space dimension enters the step size restriction (5.46) as a consequence of the definition (5.46) of $C_{\max}(h)$. Furthermore, we require $u \in C^3(J_T, H^2(\Omega))$, due to the redefinition (3.12) of $\mathcal{Y}$ and the interpolation bound (5.42). □

8.2 Example: Maxwell equations

Finally, we focus on the full discretization of the quasilinear Maxwell equations with the implicit midpoint rules (7.3) and (7.5) as well as the leapfrog scheme (7.70). To do so, we combine the space discretization from Section 5.2 with the respective abstract full discretization result Corollary 7.5 and Corollary 7.27.

Theorem 8.3. *For $p \ge 3$, let $\Omega \subset \mathbb{R}^3$ be a bounded domain with C^{p+4}-boundary. Furthermore, let the assumptions of Theorem 3.5 be satisfied for $\boldsymbol{R} \in (0, (9\|\chi\|_{L^\infty(\Omega)})^{-\frac{1}{2}})$ with $R < \boldsymbol{C}_{\boldsymbol{\mathcal{I}}}\boldsymbol{R}$ and $m = p+2$. Then, we obtain the following results for the full discretization of the discrete Maxwell equations with Kerr nonlinearity (5.57).*

(i) If there exist $\varepsilon_0, C_0 > 0$ such that the discretization parameters $h, \tau > 0$ satisfy

$$\tau < C_0 h^{\frac{5}{4}+\varepsilon_0}, \tag{8.7}$$

the full discretization of (3.7) with either of the implicit midpoint rules is well defined.

(ii) Let $\vartheta \in (0,1)$ and

$$C_1 = 2C_{\mathcal{C}}^{-1}\min\Big\{1, (1 - 9\|\chi\|_{L^\infty(\Omega)}\boldsymbol{R}^2)\frac{C_{norm}}{c_{norm}}\Big\}^{\frac{1}{2}}. \tag{8.8}$$

If there exist $\varepsilon_0, C_0 > 0$ such that the discretization parameters $h, \tau > 0$ satisfy

$$\tau < \min\Big\{C_1\vartheta h, C_0 h^{\frac{5}{4}+\varepsilon_0}\Big\}, \tag{8.9}$$

the full discretization of (3.7) with the leapfrog scheme is well defined.

In both cases, the approximations $\boldsymbol{y}_n = (\boldsymbol{\mathcal{H}}_n, \boldsymbol{\mathcal{E}}_n) \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R})$ obtained by the application of either of the schemes satisfy for $n = 0, \dots, N$

$$\|\mathcal{H}(t_n) - \boldsymbol{\mathcal{L}}_{\mathcal{V}}\boldsymbol{\mathcal{H}}_n\|_{L^2(\Omega)^3} + \|\mathcal{E}(t_n) - \boldsymbol{\mathcal{L}}_{\mathcal{H}}\boldsymbol{\mathcal{E}}_n\|_{L^2(\Omega)^3} \le C_{\mathcal{H},\mathcal{E}}(1+t_n)\mathrm{e}^{Ct_n}(h^p + \tau^2), \tag{8.10}$$

with constants $C_{\mathcal{H},\mathcal{E}}, C > 0$ independent of h, τ, and T, but $C_{\mathcal{H},\mathcal{E}}$ depending on both fields $\mathcal{H}$ and $\mathcal{E}$ as well as the nonlinear susceptibility χ, including their derivatives.

Proof. In the proof of Theorem 5.13, we already showed that Assumption 4.28 is satisfied. Furthermore, as $C_{\max}(h)$ is given by (5.73), the step size restriction (8.7) implies the restriction (7.14). Moreover, we obtain from (5.61) and (5.72) that (7.75) corresponds to

$$\tau C_{\mathcal{C}} h^{-1} \le 2\vartheta \min\Big\{1, (1 - 9\|\chi\|_{L^\infty(\Omega)}\boldsymbol{R}^2)\frac{C_{\mathrm{norm}}}{c_{\mathrm{norm}}}\Big\}^{\frac{1}{2}}.$$

Hence, (8.8) and (8.9) imply (7.93).

In the following we employ the abstract results for $y = (\mathcal{H}, \mathcal{E})$ and $\boldsymbol{y} = (\boldsymbol{\mathcal{H}}, \boldsymbol{\mathcal{E}})$, starting with the implicit midpoint rules. From Corollary 7.5, we obtain for either of the implicit midpoint rules due to $\boldsymbol{\mathcal{J}} = \boldsymbol{\mathcal{I}}$ and the definition (5.56) of the initial values the estimate

$$\begin{aligned}\|y(t_n) - \boldsymbol{\mathcal{L}}\boldsymbol{y}_n\|_{\mathcal{X}} \le \|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}})y(t_n)\|_{\mathcal{X}} + C(1+t_n)\mathrm{e}^{Ct_n}\Big(&\sup_{s_1,s_2\in[0,t_n]}\|(\mathrm{Id} - \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}})\Lambda(y(s_1))\partial_t y(s_2)\|_{\mathcal{X}} \\ &+ \sup_{s_1,s_2\in[0,t_n]} \triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\mathcal{I}}\Lambda(y(s_1))\partial_t y(s_2)) + \sup_{[0,t_n]}\|\mathcal{R}_{\mathrm{A}} y\|_{\boldsymbol{\mathcal{X}}} \\ &+ \tau^2\big(\sup_{[0,t_n]}\|\partial_t y\|_{\mathcal{Y}} + \sup_{[0,t_n]}\|\partial_t^2 y\|_{\mathcal{Y}} + \sup_{[0,t_n]}\|\partial_t^3 y\|_{\mathcal{X}}\big)\Big).\end{aligned}$$

With the bounds (5.74) and (5.75) for the remainder $\mathcal{R}_{\mathrm{A}}$ and the difference including the lift, respectively, and the interpolation property (5.62), we further get

$$\begin{aligned}&\|\mathcal{H}(t_n) - \boldsymbol{\mathcal{L}}_{\mathcal{V}}\boldsymbol{\mathcal{H}}_n\|_{L^2(\Omega)^3} + \|\mathcal{E}(t_n) - \boldsymbol{\mathcal{L}}_{\mathcal{H}}\boldsymbol{\mathcal{E}}_n\|_{L^2(\Omega)^3} \\ &\quad\le C(1+t_n)\mathrm{e}^{Ct_n}\Big(h^p\big(\sup_{[0,t_n]}\|\mathcal{H}\|_{H^{p+1}(\Omega)^3} + \sup_{[0,t_n]}\|\mathcal{E}\|_{H^{p+1}(\Omega)^3} \\ &\qquad + \sup_{s_1,s_2\in[0,t_n]}\|\Lambda(y(s_1))\partial_t y(s_2)\|_{H^p(\Omega)^3\times H^p(\Omega)^3} \\ &\qquad + \tau^2\big(\sup_{[0,t_n]}\|\partial_t\mathcal{H}\|_{H^2(\Omega)^3} + \sup_{[0,t_n]}\|\partial_t^2\mathcal{H}\|_{H^2(\Omega)^3} + \sup_{[0,t_n]}\|\partial_t^3\mathcal{H}\|_{L^2(\Omega)^3} \\ &\qquad + \sup_{[0,t_n]}\|\partial_t\mathcal{E}\|_{H^2(\Omega)^3} + \sup_{[0,t_n]}\|\partial_t^2\mathcal{E}\|_{H^2(\Omega)^3} + \sup_{[0,t_n]}\|\partial_t^3\mathcal{E}\|_{L^2(\Omega)^3}\big)\Big).\end{aligned}$$

Hence, the bound (5.76) for the nonlinearity implies (8.10) for the implicit midpoint rules.

For the leapfrog scheme, Corollary 7.27 is applicable. As most of the terms in (7.104) coincide with the terms for the implicit midpoint rules, we focus on the derivation of a bound for the additional terms in C_{τ^2}. First, we obtain as in (5.74) from Lemma 5.12 and (5.67) together with Lemma 5.11 for $\xi \in \mathcal{Y}$ the bound

$$\|\mathcal{R}_{\mathrm{A}}\xi\|_{\boldsymbol{\mathcal{X}}} \le C\|\xi\|_{\mathcal{Y}}.$$

Thus, we deduce from the definition (7.88) of $\mathrm{P}_{\mathcal{H}}$

$$\sup_{[0,t_n]} \|\mathcal{R}_{\mathrm{A}}\, \mathrm{P}_{\mathcal{H}}\, \partial_t^2 y\|_{\boldsymbol{\mathcal{X}}} \leq C \sup_{[0,t_n]} \|\partial_t^2 \mathcal{E}\|_{H^2(\Omega)^3}.$$

Furthermore, we obtain from (7.88) and (5.67)

$$\sup_{[0,t_n]} \|\mathrm{A}\, \mathrm{P}_{\mathcal{H}}\, \partial_t^2 y\|_{\mathcal{X}} \leq C \sup_{[0,t_n]} \|\partial_t^2 \mathcal{E}\|_{H^1(\Omega)^3}.$$

Finally, (7.104) together with the estimate (5.76) for the nonlinearity proves (8.10) for the leapfrog scheme. □

Note that Remark 5.15 also transfers to the full discretization of the Maxwell equations with Kerr nonlinearity. In particular, Theorem 8.3 is also valid for $\Omega \subset \mathbb{R}^3$ being a bounded domain with C^{p+3}-boundary.

CHAPTER 9

Numerical experiments

In this section we present numerical experiments which validate the error estimates in Theorem 8.1 and Theorem 8.2 for the Westervelt equation in one and two dimensions, respectively.

9.1 Example: Westervelt equation (1D)

In the following, we focus on the one-dimensional Westervelt equation. In the first part, we derive an implicit representation for the continuous solution of (3.7). We then investigate properties of this solution for a specific choice of parameters and initial values. Finally, we discuss the numerical results for the approximation of this solution.

9.1.1 General implicit representation

The derivation of the continuous solution of (3.7) is based on [Pototschnig et al., 2009, App. A], where nonlinear Maxwell equations are considered. In [Gerner, 2013, Sec. 5.4.1.1], these ideas are transfered to a quasilinear wave equation. We now apply this approach to the one-dimensional Westervelt equation.

We first consider (3.7) on the full space $\Omega = \mathbb{R}$. This is equivalent to the system

$$(1 - \varkappa u)\partial_t u = \partial_x v \qquad \text{on } J_T \times \Omega, \tag{9.1a}$$

$$\partial_t v = \partial_x u \qquad \text{on } J_T \times \Omega, \tag{9.1b}$$

with $J_T = [0, T]$. From (9.1a), the ansatz

$$v(t, x) = \varphi(u(t, x))u(t, x), \qquad t \in J_T,\ x \in \Omega, \tag{9.2}$$

for some $\varphi \in C^1(\mathbb{R})$ yields the relation

$$(1 - \varkappa u)\partial_t u = (\varphi'(u)u + \varphi(u))\partial_x u. \tag{9.3}$$

Inserting (9.2) into (9.1b) implies

$$(\varphi'(u)u + \varphi(u))\partial_t u = \partial_x u.$$

Combining these results, we obtain

$$(1 - \varkappa u)\partial_x u = \big(\varphi'(u)u + \varphi(u)\big)^2 \partial_x u.$$

In particular, the ansatz (9.2) only yields a non-trivial solution if φ satisfies

$$\varphi'(u)u + \varphi(u) = \pm(1 - \varkappa u)^{\frac{1}{2}}.$$

Hence, (9.3) corresponds to the nonlinear transport equation

$$\partial_t u = \pm(1 - \varkappa u)^{-\frac{1}{2}} \partial_x u. \tag{9.4}$$

Furthermore, if the initial values satisfy (9.4) for $t = 0$ with either the positive or the negative sign, the solution is implicitly given by

$$u(t, x) = \Phi\Big(x \pm \big(1 - \varkappa u(t, x)\big)^{-\frac{1}{2}} t\Big), \qquad t \in J_T,\ x \in \Omega, \tag{9.5}$$

where $\Phi \in C^1(\mathbb{R})$ is prescribed by the initial value u_0. In particular, for $(t, x) \in J_T \times \Omega$ fixed, the solution $u(t, x)$ is given as the fixed-point of

$$w = \Psi_{t,x}(w), \tag{9.6}$$

where $\Psi_{t,x} \in C^1(\mathbb{R})$ is given by

$$\Psi_{t,x}(w) := \Phi\Big(x \pm (1 - \varkappa w)^{-\frac{1}{2}} t\Big), \qquad t \in J_T,\ x \in \Omega. \tag{9.7}$$

We still have to justify the ansatz (9.2) by proving that (9.5) is wellposed for some $T > 0$. In particular, we have to ensure $J_T \neq \emptyset$. Then, the solution u of (9.5) is the unique solution of the Westervelt equation (3.7). Based on Banach's fixed-point theorem, this is done in the next section for a specific example.

9.1.2 Construction of a specific solution

We now fix concrete values for the numerical computations, i.e., we consider (3.7) with $\varkappa = 1$ and initial values

$$u_0(x) = \tfrac{1}{2} \exp\Big(-\frac{(x - 0.5)^2}{0.005}\Big), \qquad v_0 = \frac{x - 0.5}{0.0025}\big(1 - u_0(x)\big)^{-\frac{1}{2}} u_0(x), \qquad x \in \Omega. \tag{9.8}$$

Note that (9.4) is satisfied with the negative sign for these initial values. Hence, due to $\Omega = \mathbb{R}$, (9.5) implies

$$\Phi(s) = u_0(s), \qquad s \in \mathbb{R}. \tag{9.9}$$

We now justify the ansatz (9.2) for this specific example by pointwise application of Banach's fixed-point theorem to (9.6). To do so, let $t > 0$ and $x \in \Omega$ arbitrary, but fixed.

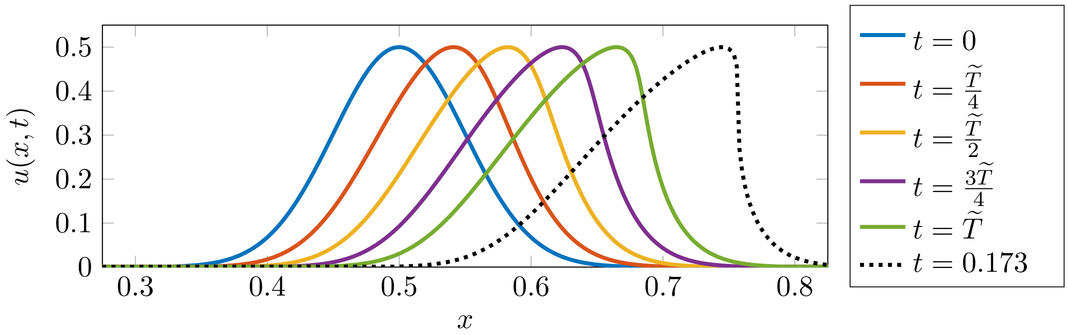

Figure 9.1: Solution u of the one-dimensional Westervelt equation (3.7) for initial values (9.8) at various points in time.

Due to $u_0(x) \in [0, \frac{1}{2}]$ $(x \in \mathbb{R})$, (9.8) and (9.9) imply

$$\Psi_{t,x} : W \to W, \qquad W = [-\tfrac{1}{2}, \tfrac{1}{2}].$$

To check that $\Psi_{t,x}$ is a contraction mapping, we obtain from (9.7) and (9.9) by the chain rule

$$\Psi'_{t,x}(w) = \Phi'\Big(x - (1-w)^{-\frac{1}{2}}t\Big)\,\frac{\partial}{\partial w}\Big(x - (1-w)^{-\frac{1}{2}}t\Big),$$
$$\Phi'(s) = \frac{0.5 - s}{0.005}\exp\left(-\frac{(s-0.5)^2}{0.005}\right).$$

Since Φ is a Gaussian function with inflection points $x_1 = 0.45$ and $x_2 = 0.55$, we have

$$\sup_{s\in\mathbb{R}} |\Phi'(s)| \leq \frac{10}{\sqrt{\mathrm{e}}}.$$

Furthermore, we get due to

$$\frac{\partial}{\partial w}\Big(x - (1-w)^{-\frac{1}{2}}t\Big) = -\tfrac{1}{2}t(1-w)^{-\frac{3}{2}}$$

the estimate

$$\sup_{w\in W}\left|\frac{\partial}{\partial w}\Big(x - (1-w)^{-\frac{1}{2}}t\Big)\right| \leq \sqrt{2}t.$$

Hence, we have shown

$$\sup_{w\in W} |\Psi'_{t,x}(w)| \leq \frac{10\sqrt{2}}{\sqrt{\mathrm{e}}}t,$$

which yields that $\Psi_{t,x}$ is a contraction mapping for $t \in J_{\widetilde{T}}$ with $\widetilde{T} = 0.1165$, since we have

$$t \leq \widetilde{T} < \frac{\sqrt{\mathrm{e}}}{10\sqrt{2}} \approx 0.11658. \tag{9.10}$$

Thus, Banach's fixed-point theorem yields the existence of a unique solution

$$u : J_{\widetilde{T}} \times \mathbb{R} \to W$$

of (3.7), which is implicitly given by (9.5) and (9.9). In particular, this implies

$$|u(t,x)| \leq \tfrac{1}{2}, \qquad\qquad t \in J_{\widetilde{t}},\ x \in \mathbb{R}. \tag{9.11}$$

Before we consider numerical results, we briefly investigate the solution u. To do so, we numerically solve of the fixed-point equation (9.5) on $J_{\widetilde{T}} \times (0,1)$. More precisely, to approximate $u(t,x)$ we use the fixed-point iteration

$$u_{t,x}^{k+1} = \Psi_{t,x}(u_{t,x}^{k}), \qquad k \geq 0, \tag{9.12}$$

with $u_{t,x}^0 = 0$ and stopping criterion

$$|u_{t,x}^{k+1} - u_{t,x}^{k}| \leq 10^{-12}.$$

Moreover, we use central difference quotients with step size 10^{-7} to approximate derivatives of u in space and time, where we again employ (9.12) to approximate the required point evaluations of u.

We point out that considering only a bounded space interval is no restriction here, as (9.5) is given locally. Furthermore, (9.11) implies

$$|x - (1-u(t,x))^{-\frac{1}{2}} t - 0.5| \geq 0.5 - \sqrt{2}\widetilde{T}, \qquad\qquad t \in J_{\widetilde{T}},\ x \in \mathbb{R}\backslash(0,1).$$

Thus, we obtain from (9.5), (9.8), and (9.9)

$$|u(t,x)| \leq \tfrac{1}{2} \exp\left(-\frac{(0.5-\sqrt{2}\widetilde{T})^2}{0.005}\right) \leq 10^{-10}, \qquad\qquad t \in J_{\widetilde{T}},\ x \in \mathbb{R}\backslash(0,1). \tag{9.13}$$

Hence, considering only the space interval $(0,1)$ is sufficient to capture essential properties of the solution u.

In Figure 9.1, a snapshot of the solution u is depicted for $t = 0, \frac{\widetilde{T}}{4}, \frac{\widetilde{T}}{2}, \frac{3\widetilde{T}}{4}, \widetilde{T}$. Additionally, despite the fact that this is not covered by our analysis with Banach's fixed-point theorem above, the solution u is shown for $t = 0.173$. We observe that the initial Gaussian function steepens in time. This self-steepening is further illustrated in Figure 9.2, where several norms of u are shown over time. Eventually, this leads to blow-up in the H^1-norm of u.

Note that a similar behavior is observed in [Pototschnig et al., 2009, App. A] for nonlinear Maxwell equations.

9.1.3 Numerical results

We now validate the error estimates from Theorem 8.1 for the full discretization of the one-dimensional Westervelt equation. In particular, we investigate the order of convergence both in space and time.

For the numerical experiments, we introduce the computational domain $\Omega = (0,1)$ and consider (3.7) subject to homogeneous Dirichlet boundary conditions. As illustrated in Figure 9.1, due to (9.13) this is a reasonable approximation. We then compute the maximal error over time, i.e.,

$$\mathrm{err} = \max_{n \leq N} \Big\{ |u(t_n) - \boldsymbol{u}_n|_{H^1(\Omega)} + \|\partial_t u(t_n) - \boldsymbol{v}_n\|_{L^2(\Omega)} \Big\}, \tag{9.14}$$

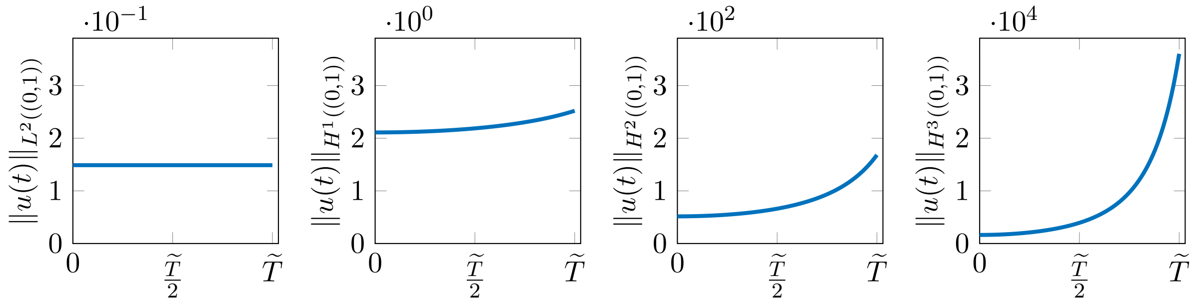

Figure 9.2: Various norms of u for initial values (9.8) over time, with $\Omega = (0, 1)$. Note the different scaling for each plot.

where we again obtain $u(t_n)$ and $v(t_n)$ with the fixed-point iteration (9.12) and the corresponding approximation by solving the fixed-point equation (9.5) numerically.

Our implementation for the one-dimensional Westervelt equation is based on the C++ finite element library `deal.II`, cf. [Bangerth et al., 2007]. For the fully implicit midpoint rule, we solve the nonlinear systems of equations using Newton's method. More precisely, based on (7.3) and (7.4) we approximate $\boldsymbol{y}_{n+1/2} = (\boldsymbol{u}_{n+1/2}, \boldsymbol{v}_{n+1/2})$ with the fixed-point iteration

$$\widehat{\boldsymbol{y}}^{k+1}_{n+1/2} = \widehat{\boldsymbol{y}}^{k}_{n+1/2} - \Big(\mathrm{D}\,\boldsymbol{G}^n\big(\widehat{\boldsymbol{y}}^{k}_{n+1/2}\big)\Big)^{-1}\boldsymbol{G}^n\big(\widehat{\boldsymbol{y}}^{k}_{n+1/2}\big), \qquad k \geq 0, \tag{9.15}$$

and $\widehat{\boldsymbol{y}}^{0}_{n+1/2} = \boldsymbol{y}_n$, where $\boldsymbol{G}^n \colon B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \to \boldsymbol{\mathcal{X}}$ is for $n = 0, \ldots, N-1$ given by

$$\boldsymbol{G}^n(\boldsymbol{\xi}) = \boldsymbol{y}_n + \tfrac{\tau}{2}\boldsymbol{\mathcal{A}}(\boldsymbol{\xi})\boldsymbol{\xi} + \tfrac{\tau}{2}\boldsymbol{\mathcal{F}}(\boldsymbol{\xi}) - \boldsymbol{\xi}, \qquad \boldsymbol{\xi} \in \boldsymbol{\mathcal{X}}.$$

As stopping criterion we use

$$\|\widehat{\boldsymbol{y}}^{k+1}_{n+1/2} - \widehat{\boldsymbol{y}}^{k}_{n+1/2}\|_{L^\infty(\Omega)} < 10^{-10}.$$

We point out that our implementation is not optimized with respect to efficiency, since we only aim at illustrating the theoretical findings. In particular, for an efficient implementation we suggest to use the simplified Newton's method. Moreover, all arising linear systems of equations are solved with the direct solver UMFPACK from [Davis, 2004], for which an interface is implemented in `deal.II`. Further details can be taken from our code, which is contained in the repository of CRC 1173 (`www.waves.kit.edu`).

On the convergence in space

In the first experiment, we investigate the dependency of the error (9.14) on the space discretization parameter h. To do so, we apply the implicit midpoint rules as well as the leapfrog scheme for a fixed number of time steps $N = 100 \cdot 2^7$, which corresponds to the time-step size $\tau \approx 10^{-5}$. We then vary the space discretization parameter

$$h \in \{2^{-k} \mid k = 1, \ldots, 18\}.$$

In Figure 9.3, we show the computational results. For the approximation space, we use the polynomial degrees $p = 2$ (solid) and $p = 3$ (dotted). Corresponding to the error estimate (8.4),

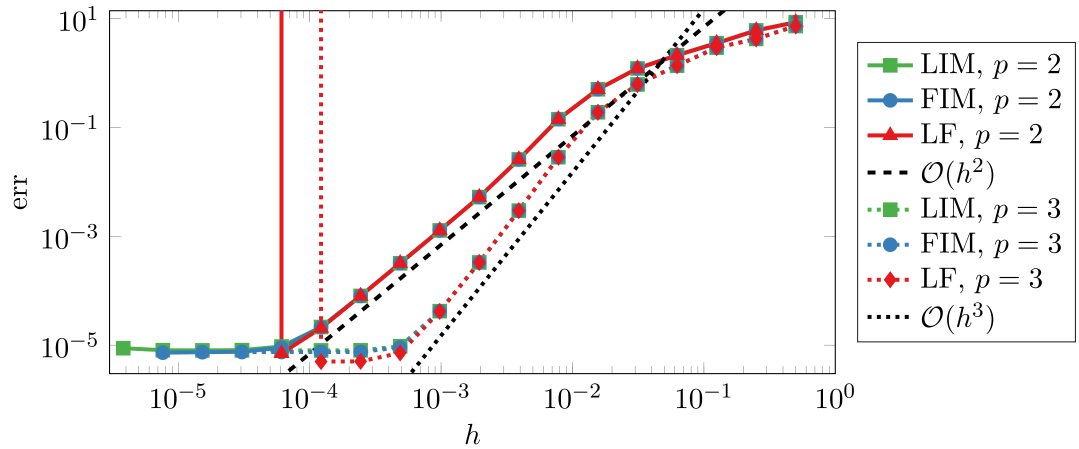

Figure 9.3: Maximal error (9.14) over time for various mesh sizes h and fixed number of time steps $N = 100 \cdot 2^7$, computed for the linearly implicit midpoint rule (LIM, green), the fully implicit midpoint rule (FIM, blue), and the leapfrog scheme (LF, red), for the space discretization with finite elements and polynomial degree $p = 2$ and $p = 3$, respectively.

for h sufficiently small there exists a regime where the error converges with order p. In this regime, all time-integration schemes yield similar results. However, this changes if we further decrease h. For the implicit midpoint rules, the time discretization error then dominates, while the leapfrog scheme becomes unstable, since the step size restriction (8.3) is no longer satisfied.

On the convergence in time

In the second experiment, we investigate the dependency of the error (9.14) on the time discretization parameter τ. Thus, we fix the mesh width $h = 2^{-12}$ and use $p = 2$ for the construction of the approximation space. We then vary the number of time steps

$$N \in \{100 \cdot 2^k \mid k = 0, \dots, 8\},$$

with corresponding time-step sizes $\tau = \frac{\widetilde{T}}{N}$.

The computational results are illustrated in Figure 9.4. For $\tau \lesssim 4 \cdot 10^{-5}$, the space discretization error dominates. Nevertheless, corresponding to the error estimate (8.4) we obtain quadratic convergence for the implicit midpoint rules for $\tau \gtrsim 4 \cdot 10^{-05}$. For the leapfrog scheme, the step size restriction (8.3) is not satisfied until the overall error is dominated by the space discretization error.

On the restrictions on the time step

Finally, we comment on the restrictions on the time step. In the numerical experiments presented so far we do not observe the step size restriction (8.1) for the implicit midpoint rules. For the leapfrog scheme, we only observe the contribution of (7.75) to (8.3), which corresponds to the classical CFL condition of the leapfrog scheme.

However, the numerical example can be modified to indicate the origin of the additional step size restriction (8.1). More precisely, the idea is to increase the considered time interval in order

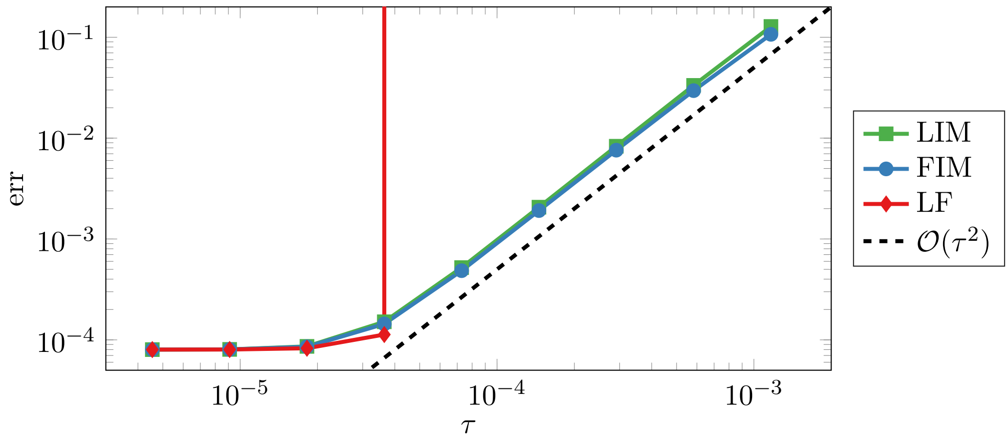

Figure 9.4: Maximal error (9.14) over time for various time-step sizes τ and fixed mesh width $h = 2^{-12}$, computed for the linearly implicit midpoint rule (LIM, green), the fully implicit midpoint rule (FIM, blue), and the leapfrog scheme (LF, red), for the space discretization with finite elements and polynomial degree $p = 2$.

to approach the blow-up of the continuous solution at the maximal time of existence $t^*(y_0)$. In fact, although Banach's fixed-point theorem only yields the upper bound $\widetilde{T} = 0.1165$ in (9.10), a numerical investigation of the fixed-point equation (9.5) implies $t^*(y_0) > 0.173$. Thus, we consider for this numerical experiment the time interval J_T with $T = 0.173$.

In Figure 9.1 the numerical approximation of the continuous solution u at $t = T$ is depicted (black, dotted). In particular, note that the slope of the wave front increased significantly compared to $t = \widetilde{T}$. Correspondingly, for the numerical approximation obtained with the linearly implicit midpoint rule, a snapshot of the front of the wave crest at time $t = T$ is depicted in Figure 9.5 for various discretization parameters. More precisely, in every row we fix the number of time steps N. Conversely, we fix the space discretization parameter h in every column. Overall, we consider

$$N \in \{100 \cdot 2^k \mid k = 3, \dots, 7\}, \qquad h \in \{2^{-k} \mid k = 12, \dots, 15\}.$$

Depending on the discretization parameters we observe numerical artifacts in the form of instabilities at the front of the wave crest. For $N \lesssim 1600$ these artifacts occur for all mesh widths h in the considered range. However, for $N \gtrsim 3200$ the instabilities only build up for h decreasing. Moreover, note that the magnitude of the oscillations is related to the time-step size.

The numerical artifacts are caused by the following two superimposing effects. On the one hand, since the bound (9.11) for the continuous solution u also holds for the time interval J_T, the interpolation $\mathcal{I}u$ is also pointwise bounded due to (5.7). However, as $\mathcal{I}u$ is piecewise polynomial, its slope is restricted by the accuracy of the space discretization, which also reflected by the inverse estimate (4.1). These arguments also transfer to the numerical solution $\boldsymbol{u}$. On the other hand, in order to approximate the time evolution of the steep wave front, the time discretization has to be sufficiently accurate.

Moreover, we emphasize that $\|\boldsymbol{y}_n\|_{\boldsymbol{\mathcal{Y}}} \leq \boldsymbol{R}$ is substantial for the wellposedness of the time-discretization schemes, since essential properties of the discrete operators in Assumption 4.1

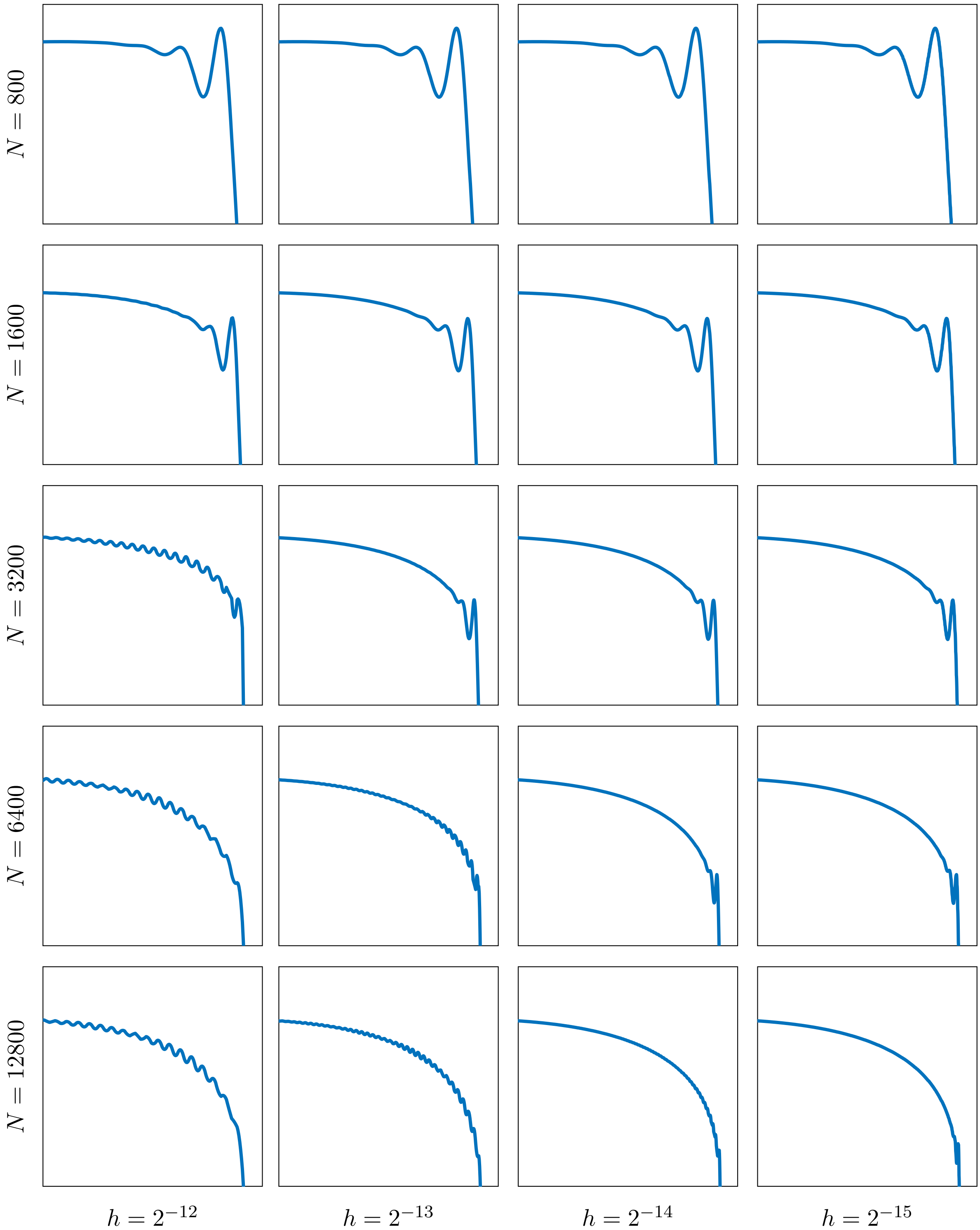

Figure 9.5: Section of the numerical solution $\boldsymbol{u}$ obtained with the linearly implicit midpoint rule, with a focus on the front of the wave crest at time $t = 0.173$. We fix the number of time steps $N = 800, 1600, 3200, 6400, 12800$ for the respective row and the space discretization parameter $h = 2^{-12}, 2^{-13}, 2^{-14}, 2^{-15}$ for the respective column.

only hold provided this pointwise estimate is satisfied. For the concrete setting considered in this section, this is satisfied for $\|\boldsymbol{u}_n\|_{L^\infty(\Omega)} < 1$. Thus, the time-integration schemes are only wellposed if the numerical artifacts are sufficiently small, which yields a relation between the time-step size τ and the space discretization parameter h, e.g., the step size restrictions (7.14) and (7.93) for the implicit midpoint rules and the leapfrog scheme, respectively. Hence, these restriction are inherent in the problem itself, as stated in [Makridakis, 1993].

To conclude, we emphasize that for a fixed time interval $[0, T]$ the slope of the continuous solution is bounded. Hence, in this case the effects described above can be neglected for τ and h sufficiently small. However, as it is unclear how to incorporate the existence of some $\delta > 0$ with $T < t^*(y_0) - \delta$ in the analysis, we have to assume the abstract step size restrictions (7.14) and (7.93).

9.2 Example: Westervelt equation (2D)

We now consider the Westervelt equation for a domain $\Omega \subset \mathbb{R}^2$ with smooth boundary. Since, up to our knowledge, a continuous solution can not be constructed as in the one-dimensional case, we add a time-dependent source term to the Westervelt equation. Using the space discretization with isoparametric elements introduced in Section 8.1.2, we then numerically investigate the error estimate for the full discretization from Theorem 8.2.

9.2.1 Modification of the Westervelt equation

We consider the following modified variant of (3.7) with an explicitly given right-hand side $f : J_T \times \Omega \to \mathbb{R}$:

$$\begin{cases} (1 - \varkappa u)\partial_t^2 u = \Delta\, u + \varkappa(\partial_t u)^2 + f & \text{on } J_T \times \Omega, \\ \qquad u(0) = u_0, \qquad \partial_t u(0) = v_0 & \text{on } \Omega. \end{cases} \tag{9.16}$$

Hence, as in (3.8), we obtain that the modified variant is of the form (3.1) with

$$y = \begin{pmatrix} u \\ v \end{pmatrix}, \quad \Lambda(y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 - \varkappa u \end{pmatrix}, \quad \mathrm{A} = \begin{pmatrix} 0 & 1 \\ \Delta & 0 \end{pmatrix}, \quad \mathrm{F}(y) = \begin{pmatrix} 0 \\ \varkappa v^2 + f \end{pmatrix}, \quad y_0 = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}.$$

Concerning the wellposedness of this modified variant, note that Theorem 3.3 is not directly applicable. However, with the same arguments as in the proof of [Kato, 1975, Thm. 6], the corresponding proof of [Dörfler et al., 2016, Thm. 4.1] can be extended to allow for additional right-hand sides. In particular, this approach yields wellposedness of (9.16) for

$$f \in C(J_T, H^2(\Omega) \cap H_0^1(\Omega)). \tag{9.17}$$

Furthermore, we emphasize that the analysis for both the space discretization in Chapter 5 and the full discretization in Chapter 8 is also valid for the modified problem if f is sufficiently smooth. In particular, we obtain that the discrete modified Westervelt equation

$$\begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & \boldsymbol{\Lambda}_{\mathcal{H}}(\boldsymbol{u}) \end{pmatrix} \begin{pmatrix} \partial_t \boldsymbol{u} \\ \partial_t \boldsymbol{v} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \boldsymbol{\Delta} & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix} + \begin{pmatrix} 0 \\ \mathbf{F}_{\mathcal{H}}(\cdot, \boldsymbol{v}) \end{pmatrix} & \text{on } J_T \times \Omega_h, \\ \qquad \boldsymbol{u}(0) = \boldsymbol{u_0}, \qquad \boldsymbol{v}(0) = \boldsymbol{v_0} & \text{on } \Omega_h, \end{cases}$$

with

$$\boldsymbol{\Lambda}_{\mathcal{H}}(\boldsymbol{\xi}_{\mathcal{V}})\boldsymbol{\zeta} = \boldsymbol{\mathcal{I}}_{\mathcal{H}}((1-\varkappa\boldsymbol{\mathcal{L}}\boldsymbol{\xi}_{\mathcal{V}})\boldsymbol{\mathcal{L}}\boldsymbol{\zeta}), \quad \mathbf{F}_{\mathcal{H}}(t,\boldsymbol{\xi}_{\mathcal{H}}) = \boldsymbol{\mathcal{I}}_{\mathcal{H}}(\varkappa(\boldsymbol{\mathcal{L}}\boldsymbol{\xi}_{\mathcal{H}})^2 + f(t)), \quad t \in J_T,\ \boldsymbol{\zeta} \in \boldsymbol{\mathcal{X}}_{\mathcal{H}},$$

is wellposed.

9.2.2 Numerical results

In this section, we investigate the order of convergence both in space and time.

For the numerical experiments, we consider the modified Westervelt equation (9.16) on the unit disc, i.e., we set $d = 2$ and $\Omega = B_{\mathbb{R}^2}(1)$. Furthermore, we set $\varkappa = 1$ and $T = 1$. Finally, the right-hand side f and the initial values u_0, v_0 are chosen such that

$$u(t,x) = \tfrac{1}{2}\sin(\pi\|x\|_2^2)^3\cos(\tfrac{\pi}{2}t), \qquad t \in [0,1],\ x \in \Omega,$$

satisfies (9.16). Since (3.11) and (9.17) are satisfied for this choice, we then have that u is the unique solution of (9.16).

In order to accelerate the numerical computations, we avoid the application of the lift operator $\boldsymbol{\mathcal{L}}$ as well as the evaluation of the error (8.6) on the exact domain Ω. Instead, we only consider for $y = (u,v)$ and $\boldsymbol{y}_n = (\boldsymbol{u}_n, \boldsymbol{v}_n)$ the maximum over time

$$\widetilde{\mathrm{err}} = \max_{n \le N}\{\|\widetilde{\boldsymbol{e}}_n\|_{\boldsymbol{\mathcal{X}}}\}, \tag{9.18}$$

for the discrete errors $\widetilde{\boldsymbol{e}}_n = \boldsymbol{\mathcal{I}}y(t_n) - \boldsymbol{y}_n$. Note that this is sufficient to validate (8.6), as the boundedness (5.35) of the lift operator $\boldsymbol{\mathcal{L}}$ implies

$$\|y(t_n) - \boldsymbol{\mathcal{L}}\boldsymbol{y}_n\|_{\mathcal{X}} \le \boldsymbol{C}_{\boldsymbol{\mathcal{L}}}\|\widetilde{\boldsymbol{e}}_n\|_{\boldsymbol{\mathcal{X}}} + \|(\boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{I}} - \mathrm{Id})y(t_n)\|_{\mathcal{X}}, \qquad n = 0,\dots,N,$$

where the second term only depends on the approximation space and the regularity of the exact solution, but is independent of the time-integration scheme.

Our implementation for the two-dimensional Westervelt equation is based on the C++ finite element library MFEM, cf. [MFEM, 2018]. Again, we solve the nonlinear schemes arising in the fully implicit midpoint rule using Newton's method (9.15) with stopping criterion

$$\|\widehat{\boldsymbol{y}}_{n+1/2}^{k+1} - \widehat{\boldsymbol{y}}_{n+1/2}^{k}\|_{L^\infty(\Omega)} < 10^{-9}.$$

The linear systems arising in either of the implicit midpoint rules are solved with the generalized minimal residual method. To accelerate the convergence, we apply a block diagonal preconditioner consisting of a Jacobi preconditioner and an algebraic multigrid preconditioner from [Falgout et al., 2006], for which an interface is implemented in MFEM. The stopping criterion is given by the relative tolerance 10^{-9}. More precisely, the iterative solution of the linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is terminated if

$$\frac{\|\boldsymbol{B}\boldsymbol{r}_m\|_2}{\|\boldsymbol{B}\boldsymbol{r}_0\|_2} \le 10^{-9}$$

is satisfied, where $\boldsymbol{B} \approx \boldsymbol{A}^{-1}$ denotes the preconditioner and $\boldsymbol{r}_m = \boldsymbol{A}\boldsymbol{x}_m - \boldsymbol{b}$ denotes the residual after $m \ge 0$ steps.

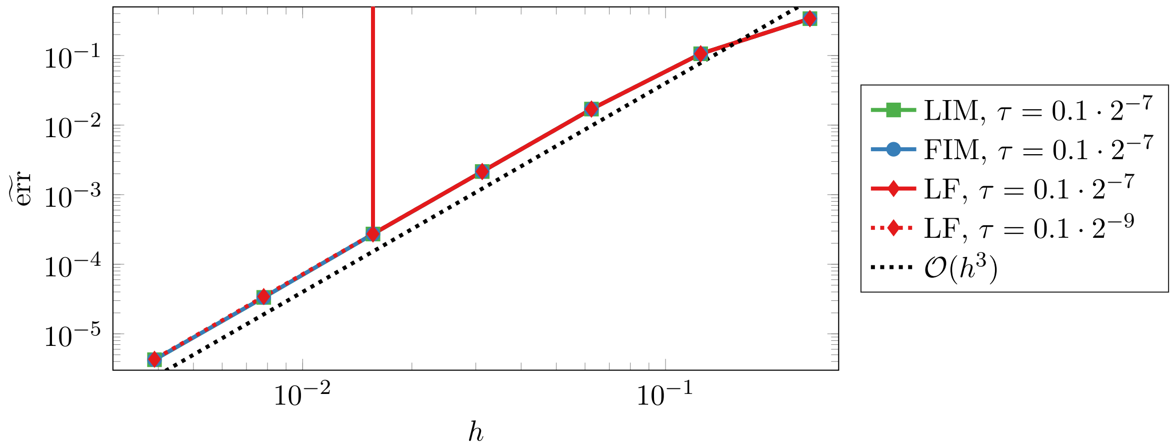

Figure 9.6: Maximal error (9.18) over time for various mesh sizes h, computed for the linearly implicit midpoint rule (LIM, green), the fully implicit midpoint rule (FIM, blue), and the leapfrog scheme (LF, red) with time-step sizes $\tau = 0.1 \cdot 2^{-7}$ and $\tau = 0.1 \cdot 2^{-9}$, respectively.

For the linear systems with the mass matrices arising in the leapfrog scheme, we use the conjugate gradient method with stopping criterion

$$\frac{\|\boldsymbol{r}_m\|_2}{\|\boldsymbol{r}_0\|_2} \leq 10^{-9},$$

where we employ the same notation as above.

Concerning the efficiency, we emphasize that the same remarks as in the one-dimensional case also apply in this case. Further details can be taken from our code, which is contained in the repository of CRC 1173 (`www.waves.kit.edu`).

On the convergence in space

The first experiment is devoted to the dependency of the error (8.6) on the space discretization parameter h. Thus, we consider the full discretization of the modified Westervelt equation (9.16) with the implicit midpoint rules and the leapfrog scheme for a fixed time-step size $\tau = 0.1 \cdot 2^{-7}$. For the space discretization, we choose $p = 3$ and

$$h \in \{2^{-k} \mid k = 2, \ldots, 8\}.$$

As illustrated in Figure 9.6, the results are almost identical for all time-integration schemes, as long as the step size restriction (8.5) for the leapfrog scheme is satisfied. In particular, we obtain cubic convergence, which corresponds to the error estimate (8.6) with polynomial degree $p = 3$. Furthermore, the application of the leapfrog scheme with reduced time-step size $\tau = 0.1 \cdot 2^{-9}$ (red, dotted) is stable for all space discretization parameters h in the range considered.

We emphasize that these results also fit well to the error estimate (5.45) for the space discretization of the Westervelt equation.

On the convergence in time

For the dependency of the error (8.6) on the time discretization parameter τ, we proceed as in the one-dimensional setting, i.e., we fix the mesh width $h = 2^{-8}$ and vary the time-step size τ.

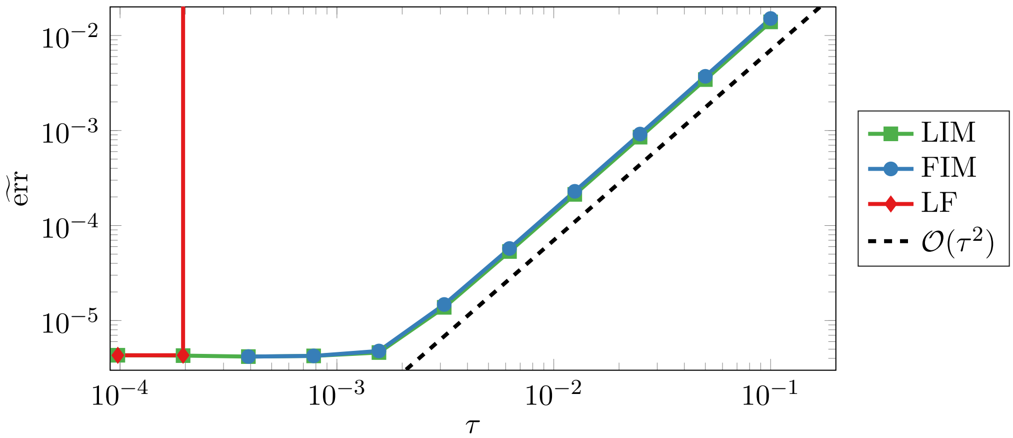

Figure 9.7: Maximal error (9.18) over time for various time-step sizes τ and fixed mesh width $h = 2^{-8}$, computed for the linearly implicit midpoint rule (LIM, green), the fully implicit midpoint rule (FIM, blue), and the leapfrog scheme (LF, red).

In particular, for the implicit midpoint rules, we consider

$$\tau \in \{0.1 \cdot 2^{-k} \mid k = 0, \dots, 7\}.$$

Since the step size restriction (8.5) of the leapfrog scheme is not satisfied for these time-step sizes, we consider

$$\tau \in \{0.1 \cdot 2^{-k} \mid k = 9, 10\}$$

for the leapfrog scheme.

In Figure 9.7 the computational results for the different time-integration schemes are depicted. For $\tau \lesssim 2 \cdot 10^{-3}$, the space discretization error dominates. For $\tau \gtrsim 2 \cdot 10^{-3}$, we obtain second order convergence for the implicit midpoint rules, which corresponds to the error estimate (8.6). As in the one-dimensional setting, the step size restriction (8.5) of the leapfrog scheme is not satisfied until the space discretization error dominates.

CHAPTER 10

Conclusion and outlook

In this thesis we presented an abstract framework to analyze the space and time discretization of a very general class of quasilinear wave-type problems. In particular, this also includes the discretization of first-order quasilinear wave-type problems, which has not been analyzed so far.

For nonconforming space discretizations of these problems, we proved both wellposedness and an error estimate based on semigroup theory. Additionally, we deduced a refined error estimate for the special case of local nonlinearities. We emphasize that the consideration of nonconforming discretizations is essential here, since the wellposedness analysis for the corresponding continuous problems is in general based on severe regularity assumptions on the boundary of the domain, which are usually not satisfied in the discrete setting.

Concerning the full discretization of quasilinear wave-type problems, we first considered a linearized version of the implicit midpoint rule, where we showed wellposedness and a rigorous error estimate using energy techniques. Based on these results and a fixed-point iteration, we extended these results to the implicit midpoint rule. Moreover, also based on the analysis of the linearized scheme we further provided a rigorous error analysis for the full discretization with the leapfrog scheme.

To emphasize the relevance of the abstract framework, we applied the abstract results to two prominent examples from physics, i.e., the Westervelt equation and the Maxwell equations with Kerr nonlinearity. To do so, we first introduced the spatially discrete setting for these equations. Based on the assumptions of the abstract framework, we then obtained rigorous error estimates for the space discretization as well as the full discretization with all three time integration schemes considered. Finally, we were able to confirm the theoretical results with numerical experiments for the Westervelt equation.

A possible extension of this thesis would be the application of higher-order time-integration schemes for the full discretization of quasilinear wave-type problems. Furthermore, the abstract framework we presented can be applied to quasilinear wave-type problems with dynamic boundary conditions.

APPENDIX A

Collection of important formulas

For the sake of readability, this is a collection of the most frequently used formulas in this thesis.

$$\begin{cases} \partial_t y(t) = \mathcal{A}(y(t))y(t) + \mathcal{F}(t, y(t)), & t \in J_T, \\ y(0) = y_0 \end{cases} \tag{3.3}$$

$$\mathcal{A}(\xi) := \Lambda(\xi)^{-1}\mathrm{A}, \qquad \mathcal{F}(t,\xi) := \Lambda(\xi)^{-1}\mathrm{F}(t,\xi), \qquad t \in J_T,\ \xi \in B_{\mathcal{Y}}(R) \tag{3.4}$$

$$\frac{1}{\boldsymbol{C}_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{Y}}}(h)}\|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}} \le \|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{Y}}} \le \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}}, \qquad \boldsymbol{\xi} \in \boldsymbol{\mathcal{Y}} \tag{4.1}$$

$$\boldsymbol{c}_{\boldsymbol{\Lambda}}\|\boldsymbol{\varphi}\|_{\boldsymbol{\mathcal{X}}}^2 \le (\boldsymbol{\Lambda}(\boldsymbol{\xi})\boldsymbol{\varphi} \mid \boldsymbol{\varphi})_{\boldsymbol{\mathcal{X}}}, \qquad \|\boldsymbol{\Lambda}(\boldsymbol{\xi})\|_{\mathcal{L}(\boldsymbol{\mathcal{X}})} \le \boldsymbol{C}_{\boldsymbol{\Lambda}}, \qquad \boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}},\ \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{4.3}$$

$$\|\boldsymbol{\Lambda}(\boldsymbol{\varphi}) - \boldsymbol{\Lambda}(\boldsymbol{\psi})\|_{\mathcal{L}(\boldsymbol{\mathcal{X}})} \le \boldsymbol{L}_{\boldsymbol{\Lambda}}^{\boldsymbol{\mathcal{X}}}\|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_{\boldsymbol{\mathcal{Y}}}, \qquad \boldsymbol{\varphi}, \boldsymbol{\psi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{4.4a}$$

$$\|\boldsymbol{\Lambda}(\boldsymbol{\varphi}) - \boldsymbol{\Lambda}(\boldsymbol{\psi})\|_{\mathcal{L}(\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}})} \le \boldsymbol{L}_{\boldsymbol{\Lambda}}^{\boldsymbol{\mathcal{Y}}}\|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}}, \qquad \boldsymbol{\varphi}, \boldsymbol{\psi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{4.4b}$$

$$(\mathbf{A}\boldsymbol{\xi} \mid \boldsymbol{\xi})_{\boldsymbol{\mathcal{X}}} \le 0, \qquad \boldsymbol{\xi} \in \boldsymbol{\mathcal{X}} \tag{4.5}$$

$$\|\mathbf{F}(t,\boldsymbol{\xi})\|_{\boldsymbol{\mathcal{Y}}} \le \boldsymbol{C}_{\mathbf{F}}, \qquad t \in J_T,\ \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{4.6}$$

$$\|\mathbf{F}(t,\boldsymbol{\varphi}) - \mathbf{F}(t,\boldsymbol{\psi})\|_{\boldsymbol{\mathcal{X}}} \le \boldsymbol{L}_{\mathbf{F}}\|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}}, \qquad t \in J_T,\ \boldsymbol{\varphi}, \boldsymbol{\psi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{4.7}$$

$$(\boldsymbol{\varphi} \mid \boldsymbol{\psi})_{\boldsymbol{\Lambda}(\boldsymbol{\xi})} := (\boldsymbol{\Lambda}(\boldsymbol{\xi})\boldsymbol{\varphi} \mid \boldsymbol{\psi})_{\boldsymbol{\mathcal{X}}}, \qquad \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}),\ \boldsymbol{\varphi}, \boldsymbol{\psi} \in \boldsymbol{\mathcal{X}} \tag{4.8}$$

$$\boldsymbol{c_\Lambda}\|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}}^2 \le \|\boldsymbol{\xi}\|_{\boldsymbol{\Lambda}(\zeta)}^2 \le \boldsymbol{C_\Lambda}\|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}}^2, \qquad \boldsymbol{\xi} \in \boldsymbol{\mathcal{X}} \tag{4.10}$$

$$\|\boldsymbol{\xi}\|_{\boldsymbol{\Lambda}(\boldsymbol{z}(t))} \le (1 + C'|t-s|)\|\boldsymbol{\xi}\|_{\boldsymbol{\Lambda}(\boldsymbol{z}(s))} \le \mathrm{e}^{C'|t-s|}\|\boldsymbol{\xi}\|_{\boldsymbol{\Lambda}(\boldsymbol{z}(s))}, \qquad s,t \in J_T,\ \boldsymbol{\xi} \in \boldsymbol{\mathcal{X}} \tag{4.11}$$

$$\begin{cases} \partial_t \boldsymbol{y}(t) = \boldsymbol{\mathcal{A}}(\boldsymbol{y}(t))\boldsymbol{y}(t) + \boldsymbol{\mathcal{F}}(t, \boldsymbol{y}(t)), & t \in J_T, \\ \boldsymbol{y}(0) = \boldsymbol{y}_0, \end{cases} \tag{4.12}$$

$$\boldsymbol{\mathcal{A}}(\boldsymbol{\xi}) \coloneqq \boldsymbol{\Lambda}(\boldsymbol{\xi})^{-1}\mathbf{A}, \qquad \boldsymbol{\mathcal{F}}(t,\boldsymbol{\xi}) \coloneqq \boldsymbol{\Lambda}(\boldsymbol{\xi})^{-1}\mathbf{F}(t,\boldsymbol{\xi}), \qquad t \in J_T,\ \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{4.13}$$

$$\|(\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi}) - \boldsymbol{\mathcal{A}}(\boldsymbol{\psi}))\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}} \le \boldsymbol{L_{\mathcal{A}}}\|\boldsymbol{\mathcal{A}}(\boldsymbol{\varphi})\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{Y}}}\|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}}, \qquad \boldsymbol{\xi} \in \boldsymbol{\mathcal{X}},\ \boldsymbol{\varphi}, \boldsymbol{\psi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{4.14}$$

$$\|\boldsymbol{\mathcal{F}}(t,\boldsymbol{\varphi}) - \boldsymbol{\mathcal{F}}(t,\boldsymbol{\psi})\|_{\boldsymbol{\mathcal{X}}} \le \boldsymbol{L_{\mathcal{F}}}\|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_{\boldsymbol{\mathcal{X}}}, \qquad t \in J_T,\ \boldsymbol{\varphi}, \boldsymbol{\psi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{4.15}$$

$$\|\boldsymbol{\Lambda}(\xi)^{-1}\|_{\mathcal{L}(\boldsymbol{\mathcal{X}})} \le \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1} \tag{4.16}$$

$$\|\boldsymbol{\mathcal{J}}\|_{\mathcal{L}(\mathcal{Y},\boldsymbol{\mathcal{X}})} \le \boldsymbol{C_{\mathcal{J}}} \tag{4.17}$$

$$\|\boldsymbol{\mathcal{I}}\|_{\mathcal{L}(\mathcal{Y},\boldsymbol{\mathcal{Y}})} \le \boldsymbol{C_{\mathcal{I}}} \tag{4.18}$$

$$\|\boldsymbol{\mathcal{L}}\|_{\mathcal{L}(\boldsymbol{\mathcal{X}},\mathcal{X})} \le \boldsymbol{C_{\mathcal{L}}} \tag{4.19}$$

$$(\boldsymbol{\mathcal{L}}\boldsymbol{\varphi} \mid \psi)_{\mathcal{X}} = (\boldsymbol{\varphi} \mid \boldsymbol{\mathcal{L}}_{\mathcal{X}}^*\psi)_{\boldsymbol{\mathcal{X}}}, \qquad \boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}},\ \psi \in \mathcal{X} \tag{4.20}$$

$$(\boldsymbol{\mathcal{L}}\boldsymbol{\varphi} \mid \psi)_{\Lambda(\xi)} = (\boldsymbol{\varphi} \mid \boldsymbol{\mathcal{L}}_{\Lambda}^*[\xi]\psi)_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)}, \qquad \xi \in B_{\mathcal{Y}}(R),\ \boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}},\ \psi \in \mathcal{X} \tag{4.21}$$

$$\|\mathbf{A}\|_{\mathcal{L}(\boldsymbol{\mathcal{X}})} \le \boldsymbol{C}_{\mathbf{A}}(h) \tag{4.29}$$

$$\boldsymbol{C_{\mathcal{I}}} R < \boldsymbol{R} \tag{4.32}$$

$$\mathcal{R}_\Lambda(\xi) \coloneqq \boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_{\mathcal{X}}^*\Lambda(\xi), \qquad \xi \in B_{\mathcal{Y}}(R) \tag{4.33}$$

$$\mathcal{R}_{\mathrm{A}} \coloneqq \mathbf{A}\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_{\mathcal{X}}^*\mathrm{A} \tag{4.34}$$

$$\mathcal{R}_{\mathrm{F}}(t,\xi) \coloneqq \mathbf{F}(t,\boldsymbol{\mathcal{I}}\xi) - \boldsymbol{\mathcal{L}}_{\mathcal{X}}^*\mathrm{F}(t,\xi), \qquad t \in J_T,\ \xi \in B_{\mathcal{Y}}(R) \tag{4.35}$$

$$\|(\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_{\Lambda}^*[\xi])\zeta\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)} \le \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-\frac{1}{2}}\|\mathcal{R}_\Lambda(\xi)\zeta\|_{\boldsymbol{\mathcal{X}}}, \qquad \zeta \in \mathcal{Y},\ \xi \in B_{\mathcal{Y}}(R) \tag{4.36}$$

$$\|(\boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}\xi)\boldsymbol{\mathcal{J}} - \boldsymbol{\mathcal{L}}_\Lambda^*[\xi]\mathcal{A}(\xi))\zeta\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)} \le \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-\frac{1}{2}}\|\mathcal{R}_{\mathrm{A}}\zeta\|_{\boldsymbol{\mathcal{X}}}, \qquad \zeta \in \mathcal{Y},\ \xi \in B_{\mathcal{Y}}(R) \tag{4.37}$$

$$\|\boldsymbol{\mathcal{F}}(t, \boldsymbol{\mathcal{I}}\xi) - \boldsymbol{\mathcal{L}}_\Lambda^*[\xi]\mathcal{F}(t,\xi)\|_{\boldsymbol{\Lambda}(\boldsymbol{\mathcal{I}}\xi)} \le \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-\frac{1}{2}}\|\mathcal{R}_{\mathrm{F}}(t,\xi)\|_{\boldsymbol{\mathcal{X}}}, \qquad t \in J_T,\ \xi \in B_{\mathcal{Y}}(R) \tag{4.38}$$

$$C_{\max}(h) = \max\{1, \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h), \boldsymbol{C}_{\boldsymbol{\mathcal{Y}},\boldsymbol{\mathcal{X}}}(h)\boldsymbol{C}_{\mathbf{A}}(h)\} \tag{4.45}$$

$$\triangle_{\boldsymbol{\mathcal{X}}}^{\boldsymbol{\mathcal{L}}}(\boldsymbol{\zeta}) \coloneqq \sup_{\|\boldsymbol{\xi}\|_{\boldsymbol{\mathcal{X}}}=1}\Big((\boldsymbol{\zeta} \mid \boldsymbol{\xi})_{\boldsymbol{\mathcal{X}}} - (\boldsymbol{\mathcal{L}}\boldsymbol{\zeta} \mid \boldsymbol{\mathcal{L}}\boldsymbol{\xi})_{\mathcal{X}}\Big), \qquad \boldsymbol{\zeta}, \boldsymbol{\xi} \in \boldsymbol{\mathcal{X}} \tag{4.57}$$

$$\widetilde{\mathcal{A}}_n \coloneqq \mathcal{A}(\widetilde{y}_n), \qquad \widetilde{\mathcal{F}}_n \coloneqq \mathcal{F}(t_n, \widetilde{y}_n), \qquad \widetilde{\boldsymbol{\mathcal{A}}}_n \coloneqq \boldsymbol{\mathcal{A}}(\boldsymbol{\mathcal{I}}\widetilde{y}_n), \qquad \widetilde{\boldsymbol{\mathcal{F}}}_n \coloneqq \boldsymbol{\mathcal{F}}(t_n, \boldsymbol{\mathcal{I}}\widetilde{y}_n) \tag{7.1}$$

$$\boldsymbol{\mathcal{A}}_n \coloneqq \boldsymbol{\mathcal{A}}(\boldsymbol{y}_n), \qquad \boldsymbol{\mathcal{F}}_n \coloneqq \boldsymbol{\mathcal{F}}(t_n, \boldsymbol{y}_n) \tag{7.2}$$

$$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + \tau\boldsymbol{\mathcal{A}}_{n+1/2}\boldsymbol{y}_{n+1/2} + \tau\boldsymbol{\mathcal{F}}_{n+1/2} \tag{7.3}$$

$$\boldsymbol{y}_{n+1/2} = \frac{\boldsymbol{y}_{n+1} + \boldsymbol{y}_n}{2} \tag{7.4}$$

$$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + \tau\underline{\boldsymbol{\mathcal{A}}}_{n+1/2}\boldsymbol{y}_{n+1/2} + \tau\underline{\boldsymbol{\mathcal{F}}}_{n+1/2} \tag{7.5}$$

$$\underline{\boldsymbol{y}}_{n+1/2} = \frac{3\boldsymbol{y}_n - \boldsymbol{y}_{n-1}}{2}, \qquad n = 1, \dots, N-1 \tag{7.6}$$

$$\widetilde{y}_{n+1} = \widetilde{y}_n + \tau\widetilde{\mathcal{A}}_{n+1/2}\widetilde{y}_{n+1/2} + \tau\widetilde{\mathcal{F}}_{n+1/2} + \delta_{n+1} \tag{7.7}$$

$$\widehat{\delta}_{n+1/2} = \widetilde{y}_{n+1/2} - \frac{\widetilde{y}_{n+1} + \widetilde{y}_n}{2} \tag{7.8}$$

$$\boldsymbol{e}_n \coloneqq \boldsymbol{\mathcal{J}}\widetilde{y}_n - \boldsymbol{y}_n, \qquad \boldsymbol{e}_{n+1/2} \coloneqq \frac{\boldsymbol{e}_{n+1} + \boldsymbol{e}_n}{2} \tag{7.9}$$

$$\boldsymbol{e}_{n+1} = \boldsymbol{e}_n + \tau\widetilde{\boldsymbol{\mathcal{A}}}_{n+1/2}\boldsymbol{e}_{n+1/2} + \tau\boldsymbol{g}_{n+1} \tag{7.10}$$

$$\tau C_{\max}(h)^{\frac{1}{2}} \le C_0 h^{\varepsilon_0} \tag{7.14}$$

$$\begin{aligned}\|\boldsymbol{g}_{\eta+1}^{\mathrm{LI}}\|_* \le C\big(&\|(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{J}})\widetilde{y}_{\eta+1/2}\|_{\boldsymbol{\mathcal{X}}} + \|\boldsymbol{\mathcal{J}}\widetilde{y}_{\eta+1/2} - \underline{\boldsymbol{y}}_{\eta+1/2}\|_* + \sup_{[t_\eta, t_{\eta+1}]}\|\mathcal{R}_\Lambda(\widetilde{y}_{\eta+1/2})\partial_t y\|_{\boldsymbol{\mathcal{X}}} \\ &+ \|\mathcal{R}_{\mathrm{A}}(\widetilde{y}_{\eta+1} + \widetilde{y}_\eta)\|_{\boldsymbol{\mathcal{X}}} + \|\mathcal{R}_{\mathrm{F}}(t_{\eta+1/2}, \widetilde{y}_{\eta+1/2})\|_{\boldsymbol{\mathcal{X}}} + \|\mathrm{A}\widehat{\delta}_{\eta+1/2}\|_{\mathcal{X}} + \|\tfrac{1}{\tau}\delta_{\eta+1}\|_{\mathcal{X}}\big)\end{aligned} \tag{7.34}$$

$$\tau\|\mathcal{J}\widetilde{y}_{1/2} - \underline{\boldsymbol{y}}_{1/2}\|^2_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{1/2})} \leq \boldsymbol{c}_{\boldsymbol{\Lambda}}^{-1}\boldsymbol{C}_{\mathcal{J}}^2 \tfrac{\tau^3}{4} \sup_{[t_0,t_{1/2}]} \|\partial_t y\|^2_{\mathcal{Y}} + (1 + C'\tfrac{\tau}{2})^2\tau\|\boldsymbol{e}_0\|^2_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_0)} \tag{7.36}$$

$$\|\mathcal{J}\widetilde{y}_{\eta+1/2} - \underline{\boldsymbol{y}}_{\eta+1/2}\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1/2})} \leq \tfrac{3}{2}\|\boldsymbol{e}_\eta\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1/2})} + \tfrac{1}{2}\|\boldsymbol{e}_{\eta-1}\|_{\boldsymbol{\Lambda}(\mathcal{I}\widetilde{y}_{\eta+1/2})} + C\tau^2 \sup_{[t_{\eta-1},t_{\eta+1/2}]} \|\partial_t^2 y\|_{\mathcal{Y}} \tag{7.39}$$

$$\boldsymbol{y}^{k+1}_{\eta+1} = \boldsymbol{y}_\eta + \tau\boldsymbol{\mathcal{A}}^k_{\eta+1/2}\boldsymbol{y}^{k+1}_{\eta+1/2} + \tau\boldsymbol{\mathcal{F}}^k_{\eta+1/2}, \qquad k \geq 0 \tag{7.42}$$

$$\boldsymbol{\mathcal{A}}^k_{\eta+1/2} \coloneqq \boldsymbol{\mathcal{A}}(\boldsymbol{y}^k_{\eta+1/2}), \qquad \boldsymbol{\mathcal{F}}^k_{\eta+1/2} \coloneqq \boldsymbol{\mathcal{F}}(t_n, \boldsymbol{y}^k_{\eta+1/2}) \tag{7.43}$$

$$\boldsymbol{y}^k_{\eta+1/2} = \frac{\boldsymbol{y}^k_{\eta+1} + \boldsymbol{y}_\eta}{2}, \qquad k \geq 0 \tag{7.44}$$

$$\|\boldsymbol{y}^k_{\eta+1}\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R} + \boldsymbol{C}_{\mathcal{I}}R), \qquad \|\boldsymbol{\mathcal{A}}^k_{\eta+1}\boldsymbol{y}^k_{\eta+1}\|_{\boldsymbol{\mathcal{Y}}} < \tfrac{1}{2}(\boldsymbol{R}^{\boldsymbol{\mathcal{A}}} + \boldsymbol{C}_{\mathcal{I}}R^{\mathcal{A}}), \qquad k \geq 0 \tag{7.50}$$

$$\mathcal{R}_-(\boldsymbol{\xi}) \coloneqq \mathrm{Id} - \tfrac{\tau}{2}\boldsymbol{\mathcal{A}}(\boldsymbol{\xi}), \qquad \mathcal{R}_+(\boldsymbol{\xi}) \coloneqq \mathrm{Id} + \tfrac{\tau}{2}\boldsymbol{\mathcal{A}}(\boldsymbol{\xi}), \qquad \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{7.58}$$

$$\mathcal{R}_-(\underline{\boldsymbol{y}}_{n+1/2})\boldsymbol{y}_{n+1} = \mathcal{R}_+(\underline{\boldsymbol{y}}_{n+1/2})\boldsymbol{y}_n + \tau\underline{\boldsymbol{\mathcal{F}}}_{n+1/2}, \qquad n = 0, \dots, N-1 \tag{7.59}$$

$$(\mathcal{R}_-(\boldsymbol{\xi})\varphi \mid \psi)_{\boldsymbol{\Lambda}(\boldsymbol{\xi})} = (\varphi \mid \mathcal{R}_+(\boldsymbol{\xi})\psi)_{\boldsymbol{\Lambda}(\boldsymbol{\xi})}, \qquad \varphi, \psi \in \boldsymbol{\mathcal{X}},\ \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{7.60a}$$

$$^2_{\boldsymbol{\Lambda}(\boldsymbol{\xi})}\,(\mathcal{R}_-(\boldsymbol{\xi})\varphi \mid \varphi)_{\boldsymbol{\Lambda}(\boldsymbol{\xi})} = (\mathcal{R}_+(\boldsymbol{\xi})\varphi \mid \varphi)_{\boldsymbol{\Lambda}(\boldsymbol{\xi})} = \|\varphi\|, \qquad \varphi \in \boldsymbol{\mathcal{X}},\ \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{7.60b}$$

$$\|\mathcal{R}_-(\boldsymbol{\xi})^{-1}\varphi\|_{\boldsymbol{\Lambda}(\boldsymbol{\xi})} \leq \|\varphi\|_{\boldsymbol{\Lambda}(\boldsymbol{\xi})}, \qquad \varphi \in \boldsymbol{\mathcal{X}},\ \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{7.60c}$$

$$\Lambda_{\mathcal{V}}(\xi) \equiv \mathrm{Id}, \qquad \boldsymbol{\Lambda}_{\boldsymbol{\mathcal{V}}}(\boldsymbol{\xi}) \equiv \mathrm{Id} \tag{7.69}$$

$$\widehat{\boldsymbol{u}}_{n+1/2} = \boldsymbol{u}_n + \tfrac{\tau}{2}\mathbf{A}_{\boldsymbol{\mathcal{V}}}\boldsymbol{v}_n + \tfrac{\tau}{2}\boldsymbol{\mathcal{F}}_{\boldsymbol{\mathcal{V}}}(t_{n+1/2}, \underline{\boldsymbol{y}}_{n+1/2}) \tag{7.70a}$$

$$\boldsymbol{v}_{n+1} = \boldsymbol{v}_n + \tau\boldsymbol{\mathcal{A}}_{\boldsymbol{\mathcal{H}}}(\underline{\boldsymbol{y}}_{n+1/2})\widehat{\boldsymbol{u}}_{n+1/2} + \tau\boldsymbol{\mathcal{F}}_{\boldsymbol{\mathcal{H}}}(t_{n+1/2}, \underline{\boldsymbol{y}}_{n+1/2}) \tag{7.70b}$$

$$\boldsymbol{u}_{n+1} = \widehat{\boldsymbol{u}}_{n+1/2} + \tfrac{\tau}{2}\mathbf{A}_{\boldsymbol{\mathcal{V}}}\boldsymbol{v}_{n+1} + \tfrac{\tau}{2}\boldsymbol{\mathcal{F}}_{\boldsymbol{\mathcal{V}}}(t_{n+1/2}, \underline{\boldsymbol{y}}_{n+1/2}) \tag{7.70c}$$

$$\mathcal{D}(\boldsymbol{\xi}) \coloneqq \begin{pmatrix} 0 & 0 \\ 0 & \tfrac{\tau}{2}\boldsymbol{\mathcal{A}}_{\boldsymbol{\mathcal{H}}}(\boldsymbol{\xi})\mathbf{A}_{\boldsymbol{\mathcal{V}}} \end{pmatrix}, \qquad \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{7.72}$$

$$\widehat{\mathcal{R}}_-(\boldsymbol{\xi}) \coloneqq \mathcal{R}_-(\boldsymbol{\xi}) + \tfrac{\tau}{2}\mathcal{D}(\boldsymbol{\xi}), \qquad \widehat{\mathcal{R}}_+(\boldsymbol{\xi}) \coloneqq \mathcal{R}_+(\boldsymbol{\xi}) + \tfrac{\tau}{2}\mathcal{D}(\boldsymbol{\xi}), \qquad \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{7.73}$$

$$\widetilde{\mathfrak{D}}_n \coloneqq \mathfrak{D}(\mathcal{I}\widetilde{y}_n), \qquad \underline{\mathfrak{D}}_n \coloneqq \mathfrak{D}(\underline{\boldsymbol{y}}_n) \tag{7.74}$$

$$\tau \boldsymbol{C}_{\mathbf{A}}(h) \le 2\vartheta \boldsymbol{c}_{\mathbf{\Lambda}} \tag{7.75}$$

$$\left(\widehat{\mathfrak{R}}_-(\boldsymbol{\xi})\boldsymbol{\varphi} \mid \boldsymbol{\psi}\right)_{\mathbf{\Lambda}(\boldsymbol{\xi})} = \left(\boldsymbol{\varphi} \mid \widehat{\mathfrak{R}}_+(\boldsymbol{\xi})\boldsymbol{\psi}\right)_{\mathbf{\Lambda}(\boldsymbol{\xi})}, \qquad \boldsymbol{\varphi}, \boldsymbol{\psi} \in \boldsymbol{\mathcal{X}},\ \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{7.76a}$$

$$\left(\widehat{\mathfrak{R}}_-(\boldsymbol{\xi})\boldsymbol{\varphi} \mid \boldsymbol{\varphi}\right)_{\mathbf{\Lambda}(\boldsymbol{\xi})} = \left(\widehat{\mathfrak{R}}_+(\boldsymbol{\xi})\boldsymbol{\varphi} \mid \boldsymbol{\varphi}\right)_{\mathbf{\Lambda}(\boldsymbol{\xi})} = \|\boldsymbol{\varphi}\|^2_{\mathbf{\Lambda}(\boldsymbol{\xi})} - \tfrac{\tau^2}{4}\|\mathbf{A}_{\boldsymbol{\mathcal{V}}}\boldsymbol{\varphi}\|^2_{\boldsymbol{\mathcal{X}}_{\mathcal{V}}}, \qquad \boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}},\ \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{7.76b}$$

$$(1-\vartheta^2)\|\boldsymbol{\varphi}\|^2_{\mathbf{\Lambda}(\boldsymbol{\xi})} \le \left(\widehat{\mathfrak{R}}_-(\boldsymbol{\xi})\boldsymbol{\varphi} \mid \boldsymbol{\varphi}\right)_{\mathbf{\Lambda}(\boldsymbol{\xi})} \le \|\boldsymbol{\varphi}\|^2_{\mathbf{\Lambda}(\boldsymbol{\xi})}, \qquad \boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}},\ \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{7.76c}$$

$$\|\widehat{\mathfrak{R}}_-(\boldsymbol{\xi})^{-1}\boldsymbol{\varphi}\|_{\mathbf{\Lambda}(\boldsymbol{\xi})} \le C_{\mathrm{stb}}\|\boldsymbol{\varphi}\|_{\mathbf{\Lambda}(\boldsymbol{\xi})}, \qquad \boldsymbol{\varphi} \in \boldsymbol{\mathcal{X}},\ \boldsymbol{\xi} \in B_{\boldsymbol{\mathcal{Y}}}(\boldsymbol{R}) \tag{7.76d}$$

$$\|\mathbf{A}_{\boldsymbol{\mathcal{V}}}\|_{\mathcal{L}(\boldsymbol{\mathcal{X}}_{\mathcal{H}},\boldsymbol{\mathcal{X}}_{\mathcal{V}})}, \|\mathbf{A}_{\boldsymbol{\mathcal{H}}}\|_{\mathcal{L}(\boldsymbol{\mathcal{X}}_{\mathcal{V}},\boldsymbol{\mathcal{X}}_{\mathcal{H}})} \le \boldsymbol{C}_{\mathbf{A}}(h) \tag{7.77}$$

$$\tau \le \min\{\boldsymbol{C}_{\mathbf{A}}(h)^{-1}\vartheta \boldsymbol{c}_{\mathbf{\Lambda}}, C_{\max}(h)^{-\frac{1}{2}}C_0 h^{\varepsilon_0}\} \tag{7.93}$$

Bibliography

[Adams and Fournier, 2003]
R. A. Adams, J. J. F. Fournier. *Sobolev spaces*, vol. 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second ed., 2003.

[Antonietti et al., 2020]
P. F. Antonietti, I. Mazzieri, M. Muhr, V. Nikolić, B. Wohlmuth. A high-order discontinuous Galerkin method for nonlinear sound waves. *Journal of Computational Physics*, 415:p. 109484, 2020. URL http://dx.doi.org/https://doi.org/10.1016/j.jcp.2020.109484.

[Bales, 1986]
L. A. Bales. Higher-order single-step fully discrete approximations for nonlinear second-order hyperbolic equations. pp. 581–604. 1986. Hyperbolic partial differential equations, III.

[Bales, 1988]
L. A. Bales. Error estimates for single step fully discrete approximations for nonlinear second order hyperbolic equations. pp. 535–553. 1988. Hyperbolic partial differential equations. V. URL http://dx.doi.org/10.1016/0898-1221(88)90279-9.

[Bales and Dougalis, 1989]
L. A. Bales, V. A. Dougalis. Cosine methods for nonlinear second-order hyperbolic equations. *Math. Comp.*, 52(186):pp. 299–319, S15–S33, 1989. URL http://dx.doi.org/10.2307/2008469.

[Bangerth et al., 2007]
W. Bangerth, R. Hartmann, G. Kanschat. deal.II – a General Purpose Object Oriented Finite Element Library. *ACM Trans. Math. Softw.*, 33(4):pp. 24/1–24/27, 2007.

[Bernardi, 1989]
C. Bernardi. Optimal finite-element interpolation on curved domains. *SIAM J. Numer. Anal.*, 26(5):pp. 1212–1240, 1989. URL http://dx.doi.org/10.1137/0726068.

[Brenner and Scott, 2008]
S. C. Brenner, L. R. Scott. *The mathematical theory of finite element methods*, vol. 15 of *Texts in Applied Mathematics*. Springer, New York, third ed., 2008. URL http://dx.doi.org/10.1007/978-0-387-75934-0.

[Busch et al., 2007]
K. Busch, G. von Freymann, S. Linden, S. Mingaleev, L. Tkeshelashvili, M. Wegener. Periodic nanostructures for photonics. *Physics Reports*, 444(3):pp. 101–202, 2007. URL http://dx.doi.org/10.1016/j.physrep.2007.02.011.

[Chen and von Wahl, 1982]
C. P. Chen, W. von Wahl. Das Rand-Anfangswertproblem für quasilineare Wellengleichungen in Sobolevräumen niedriger Ordnung. *J. Reine Angew. Math.*, 337:pp. 77–112, 1982. URL http://dx.doi.org/10.1515/crll.1982.337.77.

[Dafermos and Hrusa, 1985]
C. M. Dafermos, W. J. Hrusa. Energy methods for quasilinear hyperbolic initial-boundary value problems. Applications to elastodynamics. *Arch. Rational Mech. Anal.*, 87(3):pp. 267–292, 1985. URL http://dx.doi.org/10.1007/BF00250727.

[Davis, 2004]
T. A. Davis. Algorithm 832: UMFPACK V4.3—an unsymmetric-pattern multifrontal method. *ACM Trans. Math. Software*, 30(2):pp. 196–199, 2004. URL http://dx.doi.org/10.1145/992200.992206.

[Demlow, 2009]
A. Demlow. Higher-order finite element methods and pointwise error estimates for elliptic problems on surfaces. *SIAM J. Numer. Anal.*, 47(2):pp. 805–827, 2009. URL http://dx.doi.org/10.1137/070708135.

[Dörfler et al., 2016]
W. Dörfler, H. Gerner, R. Schnaubelt. Local well-posedness of a quasilinear wave equation. *Appl. Anal.*, 95(9):pp. 2110–2123, 2016. URL http://dx.doi.org/10.1080/00036811.2015.1089236.

[Elliott and Ranner, 2013]
C. M. Elliott, T. Ranner. Finite element analysis for a coupled bulk-surface partial differential equation. *IMA J. Numer. Anal.*, 33(2):pp. 377–402, 2013. URL http://dx.doi.org/10.1093/imanum/drs022.

[Engel and Nagel, 2000]
K.-J. Engel, R. Nagel. *One-parameter semigroups for linear evolution equations*, vol. 194 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt.

[Ern and Guermond, 2004]
A. Ern, J.-L. Guermond. *Theory and practice of finite elements*, vol. 159 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2004. URL http://dx.doi.org/10.1007/978-1-4757-4355-5.

[Evans, 2010]
L. C. Evans. *Partial differential equations*, vol. 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second ed., 2010. URL http://dx.doi.org/10.1090/gsm/019.

[Ewing, 1980]
R. E. Ewing. On efficient time-stepping methods for nonlinear partial differential equations. *Comput. Math. Appl.*, 6(1, Issu):pp. 1–13, 1980. URL http://dx.doi.org/10.1016/0898-1221(80)90055-3.

[Falgout et al., 2006]
R. D. Falgout, J. E. Jones, U. M. Yang. The design and implementation of hypre, a library of parallel high performance preconditioners. In *Numerical solution of partial differential equations on parallel computers*, vol. 51 of *Lect. Notes Comput. Sci. Eng.*, pp. 267–294. Springer, Berlin, 2006. URL http://dx.doi.org/10.1007/3-540-31619-1_8.

[Gauckler et al., 2019]
L. Gauckler, J. Lu, J. L. Marzuola, F. Rousset, K. Schratz. Trigonometric integrators for quasilinear wave equations. *Math. Comp.*, 88(316):pp. 717–749, 2019. URL http://dx.doi.org/10.1090/mcom/3339.

[Gerner, 2013]
H. Gerner. *Stetige Galerkinverfahren für zeitabhängige Maxwellgleichungen mit Kerr-Nichtlinearität*. Ph.D. thesis, 2013. URL http://dx.doi.org/10.5445/IR/1000037035.

[Hipp, 2017]
D. Hipp. *A unified error analysis for spatial discretizations of wave-type equations with applications to dynamic boundary conditions*. Ph.D. thesis, Karlsruher Institut für Technologie (KIT), 2017. URL http://dx.doi.org/10.5445/IR/1000070952.

[Hipp et al., 2019]
D. Hipp, M. Hochbruck, C. Stohrer. Unified error analysis for nonconforming space discretizations of wave-type equations. *IMA J. Numer. Anal.*, 39(3):pp. 1206–1245, 2019. URL http://dx.doi.org/10.1093/imanum/dry036.

[Hochbruck and Leibold, 2019]
M. Hochbruck, J. Leibold. Finite element discretization of semilinear acoustic wave equations with kinetic boundary conditions. CRC 1173 Preprint 2019/26, Karlsruhe Institute of Technology, 2019. URL http://dx.doi.org/10.5445/IR/1000105549.

[Hochbruck and Pažur, 2017]
M. Hochbruck, T. Pažur. Error analysis of implicit Euler methods for quasilinear hyperbolic evolution equations. *Numer. Math.*, 135(2):pp. 547–569, 2017. URL http://dx.doi.org/10.1007/s00211-016-0810-5.

[Hochbruck et al., 2018]
M. Hochbruck, T. Pažur, R. Schnaubelt. Error analysis of implicit Runge-Kutta methods for quasilinear hyperbolic evolution equations. *Numer. Math.*, 138(3):pp. 557–579, 2018. URL http://dx.doi.org/10.1007/s00211-017-0914-6.

[Hochbruck and Sturm, 2016]
M. Hochbruck, A. Sturm. Error analysis of a second-order locally implicit method for linear Maxwell's equations. *SIAM J. Numer. Anal.*, 54(5):pp. 3167–3191, 2016. URL http://dx.doi.org/10.1137/15M1038037.

[Hughes et al., 1976]
T. J. R. Hughes, T. Kato, J. E. Marsden. Well-posed quasi-linear second-order hyperbolic systems with applications to nonlinear elastodynamics and general relativity. *Arch. Rational Mech. Anal.*, 63(3):pp. 273–294 (1977), 1976. URL http://dx.doi.org/10.1007/BF00251584.

[Jacob and Zwart, 2012]
B. Jacob, H. J. Zwart. *Linear port-Hamiltonian systems on infinite-dimensional spaces*, vol. 223 of *Operator Theory: Advances and Applications*. Birkhäuser/Springer Basel AG, Basel, 2012. Linear Operators and Linear Systems. URL http://dx.doi.org/10.1007/978-3-0348-0399-1.

[Kaltenbacher and Lasiecka, 2009]
B. Kaltenbacher, I. Lasiecka. Global existence and exponential decay rates for the Westervelt equation. *Discrete and Continuous Dynamical Systems - Series S*, 2:pp. 503–523, 2009. URL http://dx.doi.org/10.3934/dcdss.2009.2.503.

[Kaltenbacher, 2015]
M. Kaltenbacher. *Numerical Simulation of Mechatronic Sensors and Actuators: Finite Elements for Computational Multiphysics*. Springer Berlin Heidelberg, 2015. URL http://dx.doi.org/10.1007/978-3-642-40170-1.

[Kato, 1970]
T. Kato. Linear evolution equations of "hyperbolic" type. *J. Fac. Sci. Univ. Tokyo Sect. I*, 17:pp. 241–258, 1970.

[Kato, 1975]
T. Kato. Quasi-linear equations of evolution, with applications to partial differential equations. In *Spectral theory and differential equations (Proc. Sympos., Dundee, 1974; dedicated to Konrad Jörgens)*, pp. 25–70. Lecture Notes in Math., Vol. 448. 1975.

[Kato, 1985]
T. Kato. *Abstract differential equations and nonlinear mixed problems*. Lezioni Fermiane. [Fermi Lectures]. Scuola Normale Superiore, Pisa; Accademia Nazionale dei Lincei, Rome, 1985.

[Koch, 1993]
H. Koch. Mixed problems for fully nonlinear hyperbolic equations. *Math. Z.*, 214(1):pp. 9–42, 1993. URL http://dx.doi.org/10.1007/BF02572388.

[Kovács and Lubich, 2018]
B. Kovács, C. Lubich. Stability and convergence of time discretizations of quasi-linear evolution equations of Kato type. *Numer. Math.*, 138(2):pp. 365–388, 2018. URL http://dx.doi.org/10.1007/s00211-017-0909-3.

[Leibold, 2017]
J. Leibold. *Semilineare Wellengleichungen mit dynamischen Randbedingungen*. Master's thesis, KIT, 2017.

[Lerch et al., 2009]
R. Lerch, G. Sessler, D. Wolf. *Technische Akustik: Grundlagen und Anwendungen*. Springer Berlin Heidelberg, 2009. URL http://dx.doi.org/10.1007/978-3-540-49833-9.

[Makridakis, 1993]
C. G. Makridakis. Finite element approximations of nonlinear elastic waves. *Math. Comp.*, 61(204):pp. 569–594, 1993. URL http://dx.doi.org/10.2307/2153241.

[Meyer and Wilke, 2011]
S. Meyer, M. Wilke. Optimal regularity and long-time behavior of solutions for the Westervelt equation. *Appl. Math. Optim.*, 64(2):pp. 257–271, 2011. URL http://dx.doi.org/10.1007/s00245-011-9138-9.

[MFEM, 2018]
MFEM. Modular finite element methods. http://mfem.org/, 2018.

[Monk, 2003]
P. Monk. *Finite element methods for Maxwell's equations*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2003. URL http://dx.doi.org/10.1093/acprof:oso/9780198508885.001.0001.

[Müller, 2014]
D. Müller. *Well-posedness for a general class of quasilinear evolution equations - with applications to Maxwell's equations*. Ph.D. thesis, KIT, 2014.

[Nikolić and Wohlmuth, 2019]
V. Nikolić, B. Wohlmuth. A priori error estimates for the finite element approximation of Westervelt's quasi-linear acoustic wave equation. *SIAM J. Numer. Anal.*, 57(4):pp. 1897–1918, 2019. URL http://dx.doi.org/10.1137/19M1240873.

[Ortner and Süli, 2007]

C. Ortner, E. Süli. Discontinuous Galerkin finite element approximation of nonlinear second-order elliptic and hyperbolic systems. *SIAM J. Numer. Anal.*, 45(4):pp. 1370–1397, 2007. URL http://dx.doi.org/10.1137/06067119X.

[Pazy, 1983]

A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, vol. 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983. URL http://dx.doi.org/10.1007/978-1-4612-5561-1.

[Pototschnig et al., 2009]

M. Pototschnig, J. Niegemann, L. Tkeshelashvili, K. Busch. Time-domain simulations of the nonlinear Maxwell equations using operator-exponential methods. *IEEE Trans. Antennas and Propagation*, 57(2):pp. 475–483, 2009. URL http://dx.doi.org/10.1109/TAP.2008.2011181.

[Schnaubelt, 2019]

R. Schnaubelt. *Lecture notes "Evolution Equations"*. WS 2018-2019. Workgroup Functional Analysis, Institute for Analysis, Karlsruhe Institute of Technology, 2018-2019. URL http://www.math.kit.edu/iana3/~schnaubelt/media/evgl-skript.pdf.

[Spitz, 2017]

M. Spitz. *Local Wellposedness of Nonlinear Maxwell Equations*. Ph.D. thesis, Karlsruher Institut für Technologie (KIT), 2017. URL http://dx.doi.org/10.5445/IR/1000078030.

[Spitz, 2019]

M. Spitz. Local wellposedness of nonlinear Maxwell equations with perfectly conducting boundary conditions. *J. Differential Equations*, 266(8):pp. 5012–5063, 2019. URL http://dx.doi.org/10.1016/j.jde.2018.10.019.

[Sturm, 2017]

A. Sturm. *Locally Implicit Time Integration for Linear Maxwell's Equations*. Ph.D. thesis, Karlsruher Institut für Technologie (KIT), 2017. URL http://dx.doi.org/10.5445/IR/1000069341.

[Westervelt, 1963]

P. J. Westervelt. Parametric Acoustic Array. *The Journal of the Acoustical Society of America*, 35(4):pp. 535–537, 1963. URL http://dx.doi.org/10.1121/1.1918525.