Karlsruhe Institute of Technology

Wave phenomena
analysis and numerics

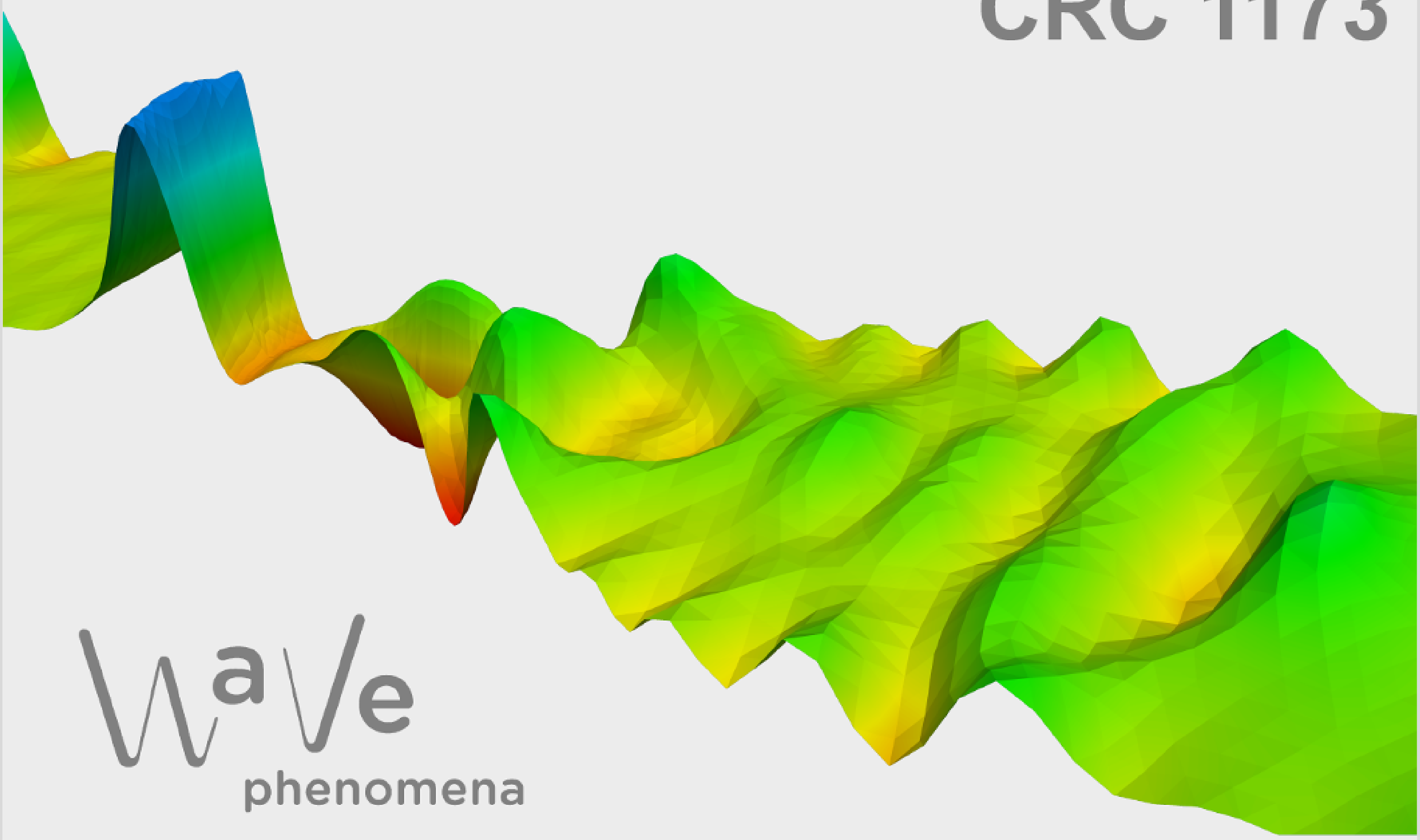
Maximum norm error bounds for the full discretization of non-autonomous wave equations

Benjamin Dörich, Jan Leibold, Bernhard Maier

CRC Preprint 2021/47, December 2021

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173



Participating universities

Universität Stuttgart

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

TU WIEN

Funded by

DFG

ISSN 2365-662X

MAXIMUM NORM ERROR BOUNDS FOR THE FULL DISCRETIZATION OF NON-AUTONOMOUS WAVE EQUATIONS

BENJAMIN DÖRICH, JAN LEIBOLD, AND BERNHARD MAIER

Abstract. In the present paper, we consider a specific class of non-autonomous wave equations on a smooth, bounded domain and their discretization in space by isoparametric finite elements and in time by the implicit Euler method. Building upon the work of Baker and Dougalis (1980), we prove maximum norm estimates for the semi discretization in space and the full discretization. The key tool is the gain of integrability coming from the inverse of the discretized differential operator. For this, we have to pay with time derivatives on the error in the L^2-norm which are reduced to estimates of the differentiated initial errors.

1. Introduction

In the present paper we consider the non-autonomous wave equation

$$\partial_{tt}u(t,x) = -\lambda(t,x)^{-1}\mathrm{L}u(t,x) + f(t,x), \qquad t \in [0,T],\ x \in \Omega, \tag{1.1}$$

with a uniformly elliptic differential operator L of order two with special emphasis on the (shifted) Laplacian. The domain $\Omega \subseteq \mathbb{R}^N$, $N = 2,3$, is assumed to be bounded and convex with a sufficiently regular boundary. On this, we impose homogeneous Dirichlet or Neumann boundary conditions. We discretize (1.1) with isoparametric finite elements in space and the implicit Euler scheme in time and derive maximum norm error bounds for the semi discretization in space and the full discretization.

A bound in the maximum norm allows us to control the numerical error at every point in the domain. Compared to the classical estimates in L^2, see, e.g., [8, 9], which are implied (with non-optimal order) by our maximum norm error estimates, and in the energy space H^1, see, e.g., [19, 30], they provide an additional insight in the approximation quality. For example, they become particularly interesting if one wants to approximate the quasilinear wave equation

$$\partial_{tt}u(t,x) = -\lambda(u(t,x))^{-1}\mathrm{L}u(t,x) + f(t,x,u(t,x)). \tag{1.2}$$

The reason is, that this equation is only well-posed as long as $\lambda(u)$ satisfies a pointwise lower bound. When discretizing (1.2) in space, it has to be ensured that the spatial discretizaion inherits this property. Since this requires a pointwise bound of the numerical approximation, maximum norm estimates, as they are

---

2020 *Mathematics Subject Classification.* Primary: 65M12, 65M15; Secondary: 65M60, 35L05.

Key words and phrases. error analysis, full discretization, wave equation, maximum norm error bounds, nonconforming space discretization, isoparametric finite elements, a-priori error bounds.

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project-ID 258734477 SFB 1173.

provided in this paper, are sufficient to guarantee such constraints. So far, an inverse inequality has to be employed, which leads to an unsatisfactory CFL condition, even for methods which are known to be unconditionally stable, or a restriction to higher-order finite elements, see, e.g., [1,30,31]. Alternatively, H^2-conforming finite elements, as suggested in [46], can be employed. For those, Sobolev's embedding can be used to obtain maximum norm estimates, once the convergence in H^2 is established. However, in order to achieve this type of conformity, the number of degrees of freedom has to be increased significantly. Our hope is to show that these constraints are only of theoretical nature and can be removed. We are confident, that the analysis presented here for the linear problem (1.1) is an important step towards the quasilinear problem (1.2) and also for higher-order methods in time.

In the articles of Baker, Dougalis, and Serbin [6,7], the space and time discretization of the linear autonomous wave-equation (i.e., $\lambda = 1, f = 0$ in (1.1)) by finite elements and one- or two-step methods, respectively, is analyzed. In our paper, we extend their analysis to the more general case of linear, non-autonomous wave equations and also to nonconforming finite elements. We point out that the latter cannot be omitted due to the following reason: In the error analysis, we rely on elliptic regularity results only available on a smooth domain Ω. Unfortunately, this prevents us from using these results on a computational domain Ω_h with a piecewise polynomial boundary. Our research is mainly inspired by [6,7] and we are not aware of further maximum norm estimates for wave equations discretized by finite elements. For finite differences on a square combined with a fourth-order in time scheme, an error bound under a CFL condition is established in [25].

For the spatial semi discretization in [6], Baker and Dougalis trade integrability, coming from the inverse of the discretized differential operator L_h, for time derivatives on the error in the L^2-norm. Those errors are controlled by the derivatives of the initial error which can be bounded using a properly preconditioned initial value. For our semi discretization, we use a similar approach and transfer the results with additional technical effort to the non-autonomous case.

For the full discretization, the proofs in [6,7] rely on an expansion of the discrete error in the eigenbasis of L_h. However, we are not aware of how to generalize this approach to the non-autonomous case. Hence, we pursue the strategy of the semi discretization. From the implicit Euler scheme we derive discrete derivatives and adapt the proofs to derive fully discrete error bounds.

In both the semi discrete and the fully discrete case, the most delicate parts are the estimates of the (discrete) derivatives of the initial errors. In the autonomous case, many terms cancel out and a spatial error bound of order $k + 1$ for order k finite elements is achieved. The additional terms arising in the non-autonomous case, however, lead to a spatial error bound of order k.

Further, we comment on maximum norm error bounds for finite element discretizations of elliptic problems as they are the fundamental tool for our error bounds in the time-dependent case. The first quasi-optimal error bounds in the maximum norm were given by Natterer [32] and Scott [43]. Many extensions and refinements have been achieved in the following years, see, e.g., [33,34,36,37,40–42, 45]. More recently in the context of nonconforming space discretizations, maximum norm error bounds for linear finite elements applied to an inhomogeneous Neumann problem were derived in [21]. For evolving surface finite element methods, similar estimates are considered in [23]. In [13], the authors of the present paper extended the approach in [11] to derive stability of a generalized Ritz map to higher-order isoparametric elements.

We also briefly comment on further work conducted in the context of maximum norm error estimates for parabolic problems. Here, we are aware of two strategies: In [5, 10], a similar approach as for the wave equation is taken and integrability is gained for time derivatives. Alternatively, some kind of stability of the semigroup generated by L_h on L^∞ is shown. This is done either directly using energy techniques, see, e.g., [38, 39], or via resolvent estimates on L^∞, see, e.g., [4, 12, 35, 44]. However, such stability estimates cannot be expected for hyperbolic problems in general, see [2, Exa. 8.4.9] and [27].

The paper is organized as follows: In Section 2, we present the analytical framework and the space discretization by isoparametric Lagrange finite elements. After providing some properties of the discretized objects, we state our main results on the error bounds for the semi discretization in space and the full discretization by the implicit Euler method.

The main parts of the proof of the semi-discrete error bound are given in Section 3. Here, we exchange the integrability in the error for time derivatives of the defect and trace those back to the initial values. We adapt the presented technique in Section 4 and transfer it from the continuous to the discrete derivatives in order to prove the theorem on the fully discrete error bound.

Section 5 is devoted to the final conclusion of our main results. We collect several approximation results and estimate the defects. Further, the (discrete) derivatives of the initial error as well as the errors in the first approximations of the fully discrete scheme are bounded.

In Appendix A, we collect some further results employed in the error analysis.

Notation. In the rest of the paper we use the notation

$$a \lesssim b\,,$$

if there is a constant $C > 0$ independent of the spatial parameter h and the time step-size τ such that $a \leq Cb$. For the sake of readability, we introduce the notation $t^n = n\tau$ and

$$x^n \coloneqq x(t^n)$$

for an arbitrary time-dependent, continuous object $x(t)$ in some Banach space X. Further, we define

$$\|x\|_{L^\infty(X)} \coloneqq \max_{[0,T]} \|x(t)\|_X\,, \qquad \|x^n\|_{\ell^\infty(X)} \coloneqq \max_{m=1,\dots,n} \|x^m\|_X\,.$$

If it is clear from the context, we write L^p instead of $L^p(\Omega)$ or $L^p(\Omega_h)$.

2. General Setting

For a convex, bounded domain $\Omega \subset \mathbb{R}^N$, $N \in \{2, 3\}$, with boundary $\partial\Omega \in C^{s,1}$, $s \in \mathbb{N}$, we study the non-autonomous wave equation (1.1) with a positive, self-adjoint operator L on $L^2(\Omega)$, stemming from a uniformly elliptic second-order differential operator with regular coefficients. Therefore, we introduce the spaces $H = L^2(\Omega)$ and $V = \mathcal{D}(\mathrm{L}^{1/2})$. The equation is further equipped with initial values

$$u(0) = u^0, \qquad\qquad \partial_t u(0) = v^0$$

and homogeneous Dirichlet or Neumann boundary conditions. Our analysis relies on the following regularity assumptions of λ.

Assumption 2.1. *There are $\kappa \in \mathbb{N}$ and $\ell_{max} \geq 1$ such that the following holds.*

(λ_1) *There exist $C_\lambda \geq c_\lambda > 0$ such that the function $\lambda\colon [0,T] \times \Omega \to \mathbb{R}$ satisfies*

$$c_\lambda \leq \lambda(t,x) \leq C_\lambda, \qquad t \in [0,T],\ x \in \Omega.$$

Moreover, we have for $\widehat{\kappa} = \max\{\kappa, \ell_{max}\}$

$$\lambda, \lambda^{-1} \in C^2\big([0,T], W^{\widehat{\kappa},\infty}(\Omega)\big) \quad \text{and} \quad \lambda \in C^3\big([0,T], L^\infty(\Omega)\big).$$

(λ_2) *For $0 \leq \ell \leq \ell_{max}$ and $u \in \mathcal{D}(\mathrm{L}^{\ell/2})$ it holds*

$$\lambda u, \lambda^{-1} u \in \mathcal{D}(\mathrm{L}^{\ell/2}).$$

We note that assumption (λ_2) guarantees that the multiplication with λ preserves the boundary conditions incorporated in L.

Example 2.2. Two admissible choices are $\mathrm{L} = -\Delta$ with Dirichlet boundary conditions ($V = H_0^1(\Omega)$) and $\mathrm{L} = -\Delta + \mathrm{Id}$ with Neumann boundary conditions ($V = H^1(\Omega)$). In this case, we have the following sufficient conditions for (λ_2).

(a) If $\nabla_x \lambda$ has compact support in Ω, (λ_2) is satisfied for any $\ell_{\max} \in \mathbb{N}$.

(b) For homogeneous Dirichlet boundary conditions, we always have $\ell_{\max} \geq 2$ and achieve $\ell_{\max} \geq 4$ by the product rule if

$$\nabla_x \lambda\big|_\Gamma = 0. \tag{2.1}$$

In the Neumann case, it holds $\ell_{\max} \geq 1$ and (2.1) yields $\ell_{\max} \geq 3$.

(c) Having the quasilinear case (1.2) in mind, and assuming $\lambda = \lambda(t,u)$ where u is the solution in V, then (2.1) follows directly in the Neumann case and in the Dirichlet case, a sufficient condition is given by $\partial_u \lambda(t,0) = 0$.

Further, condition (λ_1) directly yields the following lemma.

Lemma 2.3. *Let Assumption 2.1 be satisfied for some $\kappa \in \mathbb{N}$. Then, we have for $t \in [0,T]$, $0 \leq \ell \leq \kappa$, $1 \leq p \leq \infty$, and $j \in \{0,1,2\}$ the bounds*

$$\left\| \partial_t^j \lambda(t) \varphi \right\|_{W^{\ell,p}} \leq C_\lambda \left\| \varphi \right\|_{W^{\ell,p}}, \qquad \left\| \partial_t^j \lambda(t)^{-1} \varphi \right\|_{W^{\ell,p}} \leq C_\lambda \left\| \varphi \right\|_{W^{\ell,p}},$$

with a constant $C_\lambda > 0$ depending on λ and its derivatives.

Equivalently to (1.1), we consider the non-autonomous wave equation in first-order formulation

$$\partial_t y(t) = \Lambda(t)^{-1} \mathrm{A} y(t) + F(t), \qquad t \in [0,T], \tag{2.2}$$

with initial value $y(0) = y^0$ in the product space $X = V \times H$, with

$$y = \begin{pmatrix} u \\ \partial_t u \end{pmatrix}, \quad y^0 = \begin{pmatrix} u^0 \\ v^0 \end{pmatrix}, \quad \Lambda(t) = \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & \lambda(t) \end{pmatrix}, \quad \mathrm{A} = \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{L} & 0 \end{pmatrix}, \quad F(t) = \begin{pmatrix} 0 \\ f(t) \end{pmatrix}.$$

In particular, we emphasize that under Assumption 2.1 the operator Λ generates the time-dependent inner product

$$(\varphi \mid \psi)_{\Lambda(t)} = (\Lambda(t)\varphi \mid \psi)_X, \qquad t \in [0,T],\ \varphi, \psi \in X. \tag{2.3}$$

The corresponding norm is equivalent to the norm of X, i.e., we have

$$c_\Lambda \left\| \varphi \right\|_X^2 \leq \left\| \varphi \right\|_{\Lambda(t)}^2 \leq C_\Lambda \left\| \varphi \right\|_X^2, \qquad t \in [0,T],\ \varphi \in X, \tag{2.4}$$

with constants

$$c_\Lambda = \min\{1, c_\lambda\}, \qquad\qquad C_\Lambda = \max\{1, C_\lambda\}.$$

Further, we conclude from (λ_2) the continuity of the map

$$\Lambda(t)\colon \mathcal{D}(\mathrm{A}^\ell) \to \mathcal{D}(\mathrm{A}^\ell), \quad 0 \le \ell \le \ell_{\max},\ t \in [0,T]. \tag{2.5}$$

Our analysis relies on the solution operators of the Poisson equation in second- and first-order formulation, respectively. In particular, we introduce the second-order solution operator $S\colon H \to V$ given by

$$(S\varphi \mid \psi)_V = (\varphi \mid \psi)_H\,, \qquad\qquad \varphi \in H,\ \psi \in V. \tag{2.6}$$

For the analysis, we heavily rely on the following elliptic regularity result [17, Thm. 2.4.2.5].

Theorem 2.4 (Elliptic regularity)**.** *Let $\partial\Omega \in C^{1,1}$, then for all $1 < p < \infty$ there is a constant $C_p > 0$ such that for all $\varphi \in L^p(\Omega)$ it holds*

$$\|S\varphi\|_{W^{2,p}} \le C_p \|\varphi\|_{L^p}\,.$$

Furthermore, we define the first-order solution operator $\mathrm{T}\colon X \to \mathcal{D}(\mathrm{A})$ by

$$\mathrm{T} = \begin{pmatrix} 0 & -S \\ \mathrm{Id} & 0 \end{pmatrix}.$$

In particular, this implies $\mathrm{TA} = \mathrm{Id}$ on $\mathcal{D}(\mathrm{A})$ and $\mathrm{AT} = \mathrm{Id}$ on X.

Space discretization. We study the nonconforming space discretization of (2.2) based on isoparametric finite elements. For further details on this approach, we refer to [15]. In particular, we introduce a shape-regular and quasi-uniform mesh $\mathcal{T}_h$, consisting of isoparametric elements of degree $k \in \mathbb{N}$. The computational domain Ω_h is given by

$$\Omega_h = \bigcup_{K \in \mathcal{T}_h} K \approx \Omega,$$

where the subscript h denotes the maximal diameter of all elements $K \in \mathcal{T}_h$. Based on the transformations F_K mapping the reference element $\widehat{K}$ to $K \in \mathcal{T}_h$, we introduce the finite element space of degree k for the Neumann case

$$W_h^N = \{\varphi \in C(\overline{\Omega}) \mid \varphi|_K = \widehat{\varphi} \circ (F_K)^{-1} \text{ with } \widehat{\varphi} \in \mathcal{P}^k(\widehat{K}) \text{ for all } K \in \mathcal{T}_h\} \subset V\,,$$

and $C(\overline{\Omega})$ replaced by $C_0(\overline{\Omega})$ for W_h^D in the Dirichlet case. Here, $\mathcal{P}^k(\widehat{K})$ consists of all polynomials on $\widehat{K}$ of degree at most k. The discrete approximation spaces are given by

$$\begin{aligned} H_h^N &= \big(W_h^N, (\cdot \mid \cdot)_{L^2(\Omega_h)}\big), & H_h^D &= \big(W_h^D, (\cdot \mid \cdot)_{L^2(\Omega_h)}\big), \\ V_h^N &= \big(W_h^N, (\cdot \mid \cdot)_{H^1(\Omega_h)}\big), & V_h^D &= \big(W_h^D, (\cdot \mid \cdot)_{H^1_0(\Omega_h)}\big)\,, \end{aligned}$$

and we set $X_h = V_h \times H_h$ with $\big(V_h, H_h\big) \in \big\{\big(V_h^N, H_h^N\big), \big(V_h^D, H_h^D\big)\big\}$.

Following the detailed construction in [15, Sec. 5], we introduce the lift operator $\mathcal{L}_h\colon H_h \to H$. In particular, for $p \in [1,\infty]$ there are constants $c_p, C_p > 0$ with

$$c_p \|\varphi_h\|_{L^p(\Omega_h)} \le \|\mathcal{L}_h\varphi_h\|_{L^p(\Omega)} \le C_p \|\varphi_h\|_{L^p(\Omega_h)}\,, \qquad \varphi_h \in L^p(\Omega_h), \tag{2.7a}$$

$$c_p \|\varphi_h\|_{W^{1,p}(\Omega_h)} \le \|\mathcal{L}_h\varphi_h\|_{W^{1,p}(\Omega)} \le C_p \|\varphi_h\|_{W^{1,p}(\Omega_h)}\,, \quad \varphi_h \in W^{1,p}(\Omega_h), \tag{2.7b}$$

cf. [15, Prop. 5.8]. Further by [14, Sec. 4], the lift preserves node values, i.e. in particular

$$I_h \mathcal{L}_h \varphi_h = \varphi_h, \quad \varphi_h \in V_h \,, \tag{2.8}$$

where we denote the nodal interpolation operator by $I_h \colon C(\Omega) \to V_h$. As shown in [15, Thm. 5.9], we have for $m \in \{0,1\}$, $1 \le p \le \infty$, and $1 \le \ell \le k$ the estimates

$$\|(\mathrm{Id} - \mathcal{L}_h I_h)\varphi\|_{W^{m,p}(\Omega)} \lesssim h^{\ell+1-m} \|\varphi\|_{W^{\ell+1,p}(\Omega)} \,, \qquad \varphi \in W^{\ell+1,p}(\Omega). \tag{2.9}$$

Further, $\ell = 0$ is allowed for $N < p \le \infty$.

We define the adjoint lift operators $\mathcal{L}_h^{H*} \colon H \to H_h$ and $\mathcal{L}_h^{V*} \colon V \to V_h$ by

$$\left(\mathcal{L}_h^{H*}\varphi \mid \psi_h\right)_{H_h} = (\varphi \mid \mathcal{L}_h \psi_h)_H \,, \qquad \varphi \in H,\, \psi_h \in H_h, \tag{2.10a}$$

$$\left(\mathcal{L}_h^{V*}\varphi \mid \psi_h\right)_{V_h} = (\varphi \mid \mathcal{L}_h \psi_h)_V \,, \qquad \varphi \in V,\, \psi_h \in V_h. \tag{2.10b}$$

From [18, Thm. 5.3] and [15, Lem. 8.24], we obtain for $1 \le \ell \le k$ the bounds

$$\left\|\mathcal{L}_h^{H*}\varphi\right\|_{H_h} \lesssim \|\varphi\|_{L^2(\Omega)} \,, \qquad \varphi \in L^2(\Omega). \tag{2.11a}$$

$$\left\|(I_h - \mathcal{L}_h^{H*})\varphi\right\|_{H_h} \lesssim h^{\ell+1} \|\varphi\|_{H^{\ell+1}(\Omega)} \,, \qquad \varphi \in H^{\ell+1}(\Omega), \tag{2.11b}$$

as well as for $0 \le \ell \le k$

$$\left\|(\mathrm{Id} - \mathcal{L}_h \mathcal{L}_h^{V*})\varphi\right\|_{V_h} \lesssim h^{\ell} \|\varphi\|_{H^{\ell+1}(\Omega)} \,, \qquad \varphi \in H^{\ell+1}(\Omega). \tag{2.12}$$

For the analysis in the following sections, we additionally rely on stability and approximation properties of $\mathcal{L}_h^{V*}$. These features are well known in the literature for conforming finite elements, see, e.g., the monograph [11, Ch. 8]. In the non-conforming case, stability is shown by the authors for isoparametric finite elements in [13]. For the Neumann problem, the convergence of linear elements is studied in [21], and it will be part of our future research to extend this to higher-order isoparametric finite elements. Further, in the context of evolving surfaces similar estimates are considered in [23].

Assumption 2.5. *The adjoint lift is stable in $W^{1,\infty}$, i.e.,*

$$\left\|\mathcal{L}_h^{V*}\varphi\right\|_{W^{1,\infty}(\Omega_h)} \lesssim \|\varphi\|_{W^{1,\infty}(\Omega)} \,, \qquad \varphi \in W^{1,\infty}(\Omega).$$

For $0 \le \ell \le k$, it holds

$$\left\|(\mathrm{Id} - \mathcal{L}_h \mathcal{L}_h^{V*})\varphi\right\|_{W^{1,\infty}(\Omega)} \lesssim h^{\ell} \|\varphi\|_{W^{\ell+1,\infty}(\Omega)} \,, \qquad \varphi \in W^{\ell+1,\infty}(\Omega). \tag{2.13}$$

Combining (2.7) with (2.9) and (2.13) immediately gives the bound

$$\left\|(I_h - \mathcal{L}_h^{V*})\varphi\right\|_{W^{1,\infty}(\Omega_h)} \lesssim h^{k} \|\varphi\|_{W^{k+1,\infty}(\Omega)} \,, \qquad \varphi \in W^{k+1,\infty}(\Omega). \tag{2.14}$$

We will also employ the inverse estimate, cf. [11, Thm. 4.5.11] or [29, Lem. 5.6].

$$\|\varphi_h\|_{L^\infty} \le C h^{-N/p} \|\varphi_h\|_{L^p} \,. \tag{2.15}$$

We introduce the first-order lift operator $\boldsymbol{\mathcal{L}_h} \colon W^{\ell,p}(\Omega_h)^2 \to W^{\ell,p}(\Omega)^2$, $\ell = 0, 1$, $1 \le p \le \infty$ and reference operator $J_h \colon V \times H^2(\Omega) \to X_h$ defined by

$$\boldsymbol{\mathcal{L}_h} = \begin{pmatrix} \mathcal{L}_h & 0 \\ 0 & \mathcal{L}_h \end{pmatrix}, \qquad J_h = \begin{pmatrix} \mathcal{L}_h^{V*} & 0 \\ 0 & I_h \end{pmatrix},$$

which are bounded uniformly in h due to (2.7), (2.9) and (2.12). In particular, we have

$$J_h \in \mathcal{L}\big(\big(W^{1,\infty}(\Omega)\big)^2, \big(W^{1,\infty}(\Omega_h)\big)^2\big). \tag{2.16}$$

For $t \in [0,T]$ we define the discrete operators $\lambda_h(t) \colon H_h \to H_h$, $\Lambda_h(t) \colon X_h \to X_h$, and the discrete right-hand side $F_h(t)$ by

$$\lambda_h(t)\varphi_h = I_h\Big(\lambda(t)\mathcal{L}_h\varphi_h\Big), \quad \Lambda_h(t) = \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & \lambda_h(t) \end{pmatrix}, \quad F_h(t) = \begin{pmatrix} 0 \\ \mathcal{L}_h^{V*} f(t) \end{pmatrix}. \tag{2.17}$$

By (2.8), one can replace $\lambda(t)$ by $\mathcal{L}_h I_h \lambda(t)$. Note that if Assumption 2.1 holds for some $\kappa \geq 2$, we have $\lambda(t), f(t) \in H^2(\Omega)$ and thus by (2.7) it holds $\lambda(t)\mathcal{L}_h\varphi_h \in C(\Omega)$. Hence, the discrete operators and the discrete right-hand side are well defined. Moreover, as I_h is a nodal interpolation operator the inverse operators satisfy

$$\lambda_h(t)^{-1}\varphi_h = I_h\Big(\lambda(t)^{-1}\mathcal{L}_h\varphi_h\Big), \qquad \Lambda_h(t)^{-1} = \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & \lambda_h(t)^{-1} \end{pmatrix}. \tag{2.18}$$

Correspondingly to Lemma 2.3, we collect important properties of λ_h in the following lemma.

Lemma 2.6. *Let Assumption 2.1 be satisfied for some $\kappa \geq 2$. Then, we have for $t \in [0,T]$, $\ell \in \{0,1\}$, $1 \leq p \leq \infty$, and $j \in \{0,1,2\}$ the bounds*

$$\left\|\partial_t^j \lambda_h(t)\varphi_h\right\|_{W^{\ell,p}} \leq C_\lambda \left\|\varphi_h\right\|_{W^{\ell,p}}, \qquad \left\|\partial_t^j \lambda_h(t)^{-1}\varphi_h\right\|_{W^{\ell,p}} \leq C_\lambda \left\|\varphi_h\right\|_{W^{\ell,p}},$$

with a constant $C_\lambda > 0$ depending only on λ and its derivatives.

Proof. We note that by (2.17) and (2.18) it is sufficient to find a constant C independent of h such that for $\varphi_h, \psi_h \in V_h$ it holds

$$\left\|I_h\big(\mathcal{L}_h\varphi_h\mathcal{L}_h\psi_h\big)\right\|_{W^{\ell,p}(\Omega_h)} \leq C\left\|\varphi_h\psi_h\right\|_{W^{\ell,p}(\Omega_h)} \tag{2.19}$$

for $\ell \in \{0,1\}$, $1 \leq p \leq \infty$. This can however be reduced by the affine transformation to the reference element $\widehat{K}$, cf. [15, Lem. 4.12], where the constant then only depends on N and k, i.e., the (finite) dimension of $\mathcal{P}^k(\widehat{K})$. ☐

Furthermore, due to the definitions of J_h and Λ_h based on the nodal interpolation operator I_h, we particularly have by (2.8) the identities

$$\Lambda_h(t)J_h = J_h\Lambda(t), \qquad \Lambda_h(t)^{-1}J_h = J_h\Lambda(t)^{-1}. \tag{2.20}$$

Finally, we introduce the operators $\mathrm{L}_h \colon V_h \to H_h$ and $\mathrm{A}_h \colon X_h \to X_h$ given by

$$(\mathrm{L}_h\varphi_h \mid \psi_h)_{H_h} = (\varphi_h \mid \psi_h)_{V_h}, \qquad \mathrm{A}_h = \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{L}_h & 0 \end{pmatrix}, \qquad \varphi_h, \psi_h \in V_h.$$

Note that these operators are not uniformly bounded with respect to h.

Correspondingly to (2.3) and (2.4), the discrete operator Λ_h generates the time-dependent inner product

$$(\varphi_h \mid \psi_h)_{\Lambda_h(t)} = (\Lambda_h(t)\varphi_h \mid \psi_h)_{X_h}, \qquad t \in [0,T],\ \varphi_h, \psi_h \in X_h,$$

with the induced norm being equivalent to the norm of X_h, i.e., we have

$$c_{\Lambda_h}\left\|\varphi_h\right\|_{X_h}^2 \leq \left\|\varphi_h\right\|_{\Lambda_h(t)}^2 \leq C_{\Lambda_h}\left\|\varphi_h\right\|_{X_h}^2, \qquad t \in [0,T],\ \varphi_h \in X_h. \tag{2.21}$$

We define the discrete solution operator $S_h = \mathrm{L}_h^{-1} \colon H_h \to V_h$ by

$$(S_h\varphi_h \mid \psi_h)_{V_h} = (\varphi_h \mid \psi_h)_{H_h}, \qquad \varphi_h, \psi_h \in V_h, \tag{2.22}$$

and further the corresponding first-order solution operator

$$\mathrm{T}_h = \begin{pmatrix} 0 & -S_h \\ \mathrm{Id} & 0 \end{pmatrix},$$

which again satisfies $\mathrm{T}_h \mathrm{A}_h = \mathrm{Id}$ and $\mathrm{A}_h \mathrm{T}_h = \mathrm{Id}$.

The spatially discrete non-autonomous wave equation in first-order formulation then reads

$$\partial_t y_h(t) = \Lambda_h(t)^{-1} \mathrm{A}_h y_h(t) + F_h(t), \qquad t \in [0,T], \tag{2.23}$$

with the initial value

$$y_h(0) = y_h^0 = \mathrm{T}_h \Lambda_h(0) \mathrm{T}_h J_h \mathrm{A} \Lambda(0)^{-1} \mathrm{A} y^0. \tag{2.24}$$

Due to (2.20), this corresponds to

$$y_h^0 = \big(\mathrm{T}_h \Lambda_h(0)\big)^2 J_h \big(\Lambda(0)^{-1}\mathrm{A}\big)^2 y^0.$$

We emphasize that this choice, which is motivated by [6], is crucial for our analysis. For error bounds of differentiated errors such choices also had to be employed in [16, 22]. Similarly, the right-hand side F_h defined in (2.17) satisfies

$$F_h = \mathrm{T}_h J_h \mathrm{A} F. \tag{2.25}$$

In the spatially continuous case, the solution operator S can be used to obtain regularity which is traded in for pointwise estimates via the bounded map

$$S \colon L^2 \to H^2 \hookrightarrow L^\infty.$$

However, since we use Lagrangian finite elements which are *not* H^2-conforming, this approach does not work with S_h. Hence, in the following we provide estimates of S_h that directly give us integrability without a detour via higher-order Sobolev spaces. The following result has already been proven in [10, Lem. 4.1] in the conforming case only, and we give an adapted version of the proof in Appendix A. Since we work throughout the paper with the first-order formulation, we state the result for T_h.

Lemma 2.7. *Let Assumption 2.5 be satisfied, $\partial\Omega \in C^{1,1}$, $p \geq 2$ and $q, r \geq 1$ with $0 \leq \frac{1}{r} - \frac{1}{p} < \frac{1}{N}$. Then, the solution operator T_h satisfies*

$$\|\mathrm{T}_h \xi_h\|_{L^p \times L^q} \lesssim \|\xi_h\|_{L^q \times L^r}$$

for $\xi_h \in X_h$.

A direct consequence of the above lemma for $N = 2, 3$, is the possibility to consider the maps

$$X_h \hookrightarrow L^4 \times L^2 \xrightarrow{\mathrm{T}_h} L^4 \times L^4 \xrightarrow{\mathrm{T}_h} L^\infty \times L^4 \xrightarrow{\mathrm{T}_h} L^\infty \times L^\infty, \tag{2.26}$$

which allow us to bound the maximum norm $\|\cdot\|_{L^\infty \times L^\infty}$ in terms of the energy norm $\|\cdot\|_X$ if we apply the solution operator T_h sufficiently often. We explain in Section 3 how to employ this observation.

We can finally state our first main result on the semi discretization. The proof is given in Section 3. We use the notation

$$k^* = \max\{k, 2\}$$

in order to treat linear and higher-order finite elements simultaneously.

Theorem 2.8. *Let $\partial\Omega \in C^{k+1,1}$, let Assumption 2.1 hold for $\ell_{max} \geq 2$ and $\kappa = k+1$, and let Assumption 2.5 be satisfied. Further, assume that the solution of (1.1) satisfies*

$$\begin{aligned} y &\in C^3\big([0,T], W^{k+1,\infty}(\Omega) \times W^{k+1,N+1}(\Omega)\big) \\ &\cap C^3\big([0,T], \big(H^{k^*+2}(\Omega) \cap \mathcal{D}(\mathrm{L})\big) \times \big(H^{k^*+1}(\Omega) \cap \mathcal{D}(\mathrm{L})\big)\big) \\ &\cap C\big([0,T], W^{k+1,\infty}(\Omega) \times W^{k,\infty}(\Omega)\big)\,, \end{aligned}$$

and the right-hand side

$$f \in C^3\big([0,T], H^{k^*}(\Omega)\big) \cap C^2\big([0,T], W^{k+1,\infty}(\Omega)\big) \cap C\big([0,T], \big(H^{k^*+2}(\Omega) \cap \mathcal{D}(\mathrm{L})\big)\,,$$

and the initial value $y_h(0)$ is chosen as in (2.24). Then we have the error bound

$$\|y(t) - \boldsymbol{\mathcal{L}_h} y_h(t)\|_{L^\infty \times L^\infty} \leq C h^k,$$

where C is independent of h.

Full discretization. We study the full discretization with the backward Euler scheme

$$\partial_\tau y_h^n = (\Lambda_h^n)^{-1} \mathrm{A}_h y_h^n + F_h^n, \tag{2.27}$$

where $\tau > 0$ denotes the time step and the discrete approximation of the time derivative is for a sequence (φ^n) given by

$$\partial_\tau \varphi^n = \frac{\varphi^n - \varphi^{n-1}}{\tau}. \tag{2.28}$$

For the fully discrete scheme, we introduce the initial value

$$y_h^0 = \mathrm{T}_h \Lambda_h^1 \mathrm{T}_h J_h \mathrm{A}(\Lambda^1)^{-1} \mathrm{A} y^0 \tag{2.29}$$

which corresponds, due to (2.20), to

$$y_h^0 = \mathrm{T}_h \Lambda_h^1 \mathrm{T}_h \Lambda_h^2 J_h (\Lambda^2)^{-1} \mathrm{A}(\Lambda^1)^{-1} \mathrm{A} y^0.$$

Our second main result on the full discretization, which is proved in Section 4, then reads as follows.

Theorem 2.9. *Let $\partial\Omega \in C^{k+1,1}$, let Assumption 2.1 hold for $\ell_{max} \geq 4$ and $\kappa = k+1$, and let Assumption 2.5 be satisfied. Further, assume that the solution of (1.1) satisfies*

$$\begin{aligned} y &\in C^5([0,T], H^1(\Omega) \times H^2(\Omega)) \\ &\cap C^4([0,T], W^{1,\infty}(\Omega) \times W^{1,N+1}(\Omega)) \\ &\cap C^3([0,T], W^{k+1,\infty}(\Omega) \times W^{k+1,N+1}(\Omega)) \\ &\cap C^3\big([0,T], \big(H^{k^*+2}(\Omega) \cap \mathcal{D}(\mathrm{L})\big) \times \big(H^{k+1}(\Omega) \cap \mathcal{D}(\mathrm{L})\big)\big) \\ &\cap C^2([0,T], \mathcal{D}(\mathrm{A}^2)) \\ &\cap C([0,T], W^{k+1,\infty}(\Omega) \times W^{k,\infty}(\Omega)) \cap C([0,T], \mathcal{D}(\mathrm{A}^{k^*+3}))\,, \end{aligned}$$

and the right-hand side

$$f \in C^3\big([0,T], H^{k^*}(\Omega)\big) \cap C^2\big([0,T], W^{k+1,\infty}(\Omega) \cap \mathcal{D}(\mathrm{L}^{k^*/2})\big) \cap C\big([0,T], \mathcal{D}(\mathrm{L}^{k^*/2+1})\big)\,,$$

and the initial value y_h^0 is chosen as in (2.29). Then, there is $\tau_0 > 0$ such that for $\tau \leq \tau_0$ we have the error bound

$$\|y(t^n) - \boldsymbol{\mathcal{L}_h} y_h^n\|_{L^\infty \times L^\infty} \leq C\tau + C h^{\min\{k, \ell_{max}-2\}}, \quad n \geq 3,$$

where C is independent of h and τ, and τ_0 is independent of h.

We refer to Remark 5.9 below in order to explain the minimum in the convergence rate. Further, we emphasize that the first three approximations do not enter the above error bound. If we want to bound them as well, we need the following resolvent estimates in the maximum norm.

Assumption 2.10. *It holds the resolvent estimate for $\mu > 0$*

$$\left\|(\mu + \lambda_h(t)^{-1}\mathrm{L}_h)^{-1}x\right\|_{L^\infty(\Omega_h)} \leq C\frac{|\log(h)|}{1+|\mu|}\,\|x\|_{L^\infty(\Omega_h)}$$

with a constant C independent of h and t.

Remark 2.11. Resolvent estimates of this type have been shown in the literature at least for the case $\lambda(t,x) = \lambda_0 \in \mathbb{R}_{>0}$. For example, in [3, Thm. 2.2] and [4, Thm. 1.1] homogeneous Dirichlet boundary conditions and Lagrangian finite elements on smooth domains were considered and in certain cases the logarithmic factor can be removed. In [24, Thm. 15], the case $N = 3$ on polyhedral domains is studied. However, since the technique of proof is related to the ones in Assumption 2.5, i.e. [13], [23], we expect that this proof extends to our more general case.

This yields the convergence also of the first approximations.

Theorem 2.12. *Let the assumptions of Theorem 2.9 hold. If additionally Assumption 2.10 is valid, we obtain*

$$\left\|y(t^\ell) - \boldsymbol{\mathcal{L}_h} y_h^\ell\right\|_{L^\infty\times L^\infty} \leq C\big(\tau + h^k\big)|\log(h)|^\ell, \quad \ell = 0,1,2.$$

Extension to other $L^p \times L^q$ bounds. We briefly comment on how to extend the above stated results to dimension $N = 1$ as well as error bounds in the spaces

$$L^\infty \times L^{p^*} \quad \text{and} \quad L^{p^*} \times L^{p^*},$$

for $p^* < \frac{2N}{N-2}$, since they are easily concluded from the proofs presented in this paper. The main modifications are given by an adaption of (2.26).

In the case $N = 1$, we use Sobolev's embedding to see

$$X_h \hookrightarrow L^\infty \times L^2 \xrightarrow{\mathrm{T}_h} L^\infty \times L^\infty,$$

and the proofs shorten significantly as we only have to apply T_h once. Further, we simplify the initial value in (2.29) and the right-hand side to

$$y_h^0 = J_h y^0, \qquad F_h(t) = \begin{pmatrix} 0 \\ I_h f(t) \end{pmatrix}. \tag{2.30}$$

Similarly, for $N = 2,3$ we employ a variant of (2.26)

$$X_h \hookrightarrow L^{p^*} \times L^2 \xrightarrow{\mathrm{T}_h} L^{p^*} \times L^{p^*} \xrightarrow{\mathrm{T}_h} L^\infty \times L^{p^*}$$

such that we only need to apply T_h once or twice, respectively. For a bound in $L^{p^*} \times L^{p^*}$ we use the initial value (2.30) and for $L^\infty \times L^{p^*}$ we take

$$y_h^0 = \mathrm{T}_h J_h \mathrm{A} y^0, \qquad F_h(t) = \begin{pmatrix} 0 \\ I_h f(t) \end{pmatrix}. \tag{2.31}$$

From these, the corresponding versions of Theorems 2.8 and 2.9 are derived along the lines of the following sections.

Corollary 2.13. *Let $\partial\Omega \in C^{k+1,1}$. Further, let y be the solution of (1.1), and let the initial value $y_h(0)$ be chosen as in (2.31).*

(a) Let Assumption 2.1 hold for $\ell_{max} \geq 1$, and let y satisfy

$$\begin{aligned} y \in\ & C^2\big([0,T], W^{k+1,\infty}(\Omega) \times W^{k+1,N+1}(\Omega)\big) \\ &\cap C^2\big([0,T], \big(H^{k^*+2}(\Omega) \cap \mathcal{D}(\mathrm{L})\big) \times \big(H^{k^*+1}(\Omega) \cap \mathcal{D}(\mathrm{L})\big)\big) \\ &\cap C\big([0,T], W^{k+1,\infty}(\Omega) \times W^{k,\infty}(\Omega)\big)\,, \end{aligned}$$

and the right-hand side

$$\begin{aligned} f \in & C^2\big([0,T], H^{k^*}(\Omega)\big) \cap C^1\big([0,T], W^{k+1,\infty}(\Omega)\big) \\ &\cap C\big([0,T], H^{k^*+1}(\Omega) \cap \mathcal{D}(\mathrm{L})\big)\,. \end{aligned}$$

Then, we have the error bound

$$\|y(t) - \boldsymbol{\mathcal{L}_h} y_h(t)\|_{L^\infty \times L^{p^*}} \leq Ch^k,$$

where C is independent of h.

(b) Let Assumption 2.1 hold for $\ell_{max} \geq 3$ and let y satisfy

$$\begin{aligned} y \in\ & C^4([0,T], H^1(\Omega) \times H^2(\Omega)) \\ &\cap C^3([0,T], W^{1,\infty}(\Omega) \times W^{1,N+1}(\Omega)) \\ &\cap C^2([0,T], W^{k+1,\infty}(\Omega) \times W^{k+1,N+1}(\Omega)) \\ &\cap C^2\big([0,T], \big(H^{k^*+2}(\Omega) \cap \mathcal{D}(\mathrm{L})\big) \times \big(H^{k+1}(\Omega) \cap \mathcal{D}(\mathrm{L})\big)\big) \\ &\cap C^1([0,T], \mathcal{D}(\mathrm{A}^2)) \\ &\cap C([0,T], W^{k+1,\infty}(\Omega) \times W^{k,\infty}(\Omega)) \cap C([0,T], \mathcal{D}(\mathrm{A}^{k^*+2})) \end{aligned}$$

and the right-hand side

$$f \in C^2\big([0,T], W^{k+1,\infty}(\Omega) \cap \mathcal{D}(\mathrm{L}^{k^*/2})\big) \cap C\big([0,T], \mathcal{D}(\mathrm{L}^{k^*/2+1})\big)\,.$$

Then, we have the error bound

$$\|y(t^n) - \boldsymbol{\mathcal{L}_h} y_h^n\|_{L^\infty \times L^{p^*}} \leq C\tau + Ch^{\min\{k, \ell_{max}-1\}}, \quad n \geq 2,$$

where C is independent of h.

A bound on the first approximation can be achieved similarly to Theorem 2.12.

3. Analysis of the space discretization

3.1. **Strategy of the proof.** We now prove Theorem 2.8, i.e., we derive an error bound for the spatially discrete approximation obtained by (2.23) in the maximum norm. To this end, we proceed as follows. We split the error in

$$y(t) - \boldsymbol{\mathcal{L}_h} y_h(t) = \big(\mathrm{Id} - \boldsymbol{\mathcal{L}_h} J_h\big) y(t) + \boldsymbol{\mathcal{L}_h}\big(J_h y(t) - y_h(t)\big) =: e_{J_h}(t) + \boldsymbol{\mathcal{L}_h} e_h(t) \tag{3.1}$$

and derive an equation for the discrete error e_h. With the solution operator T, we rewrite (2.2) as

$$\mathrm{T}\Lambda(t)\partial_t y = y + \mathrm{T}\Lambda(t)F(t), \qquad t \in [0,T],$$

with initial value $y(0) = y^0$. Correspondingly, we use the discrete solution operator T_h to obtain from (2.23) the semi-discrete equation

$$\mathrm{T}_h\Lambda_h(t)\partial_t y_h = y_h + \mathrm{T}_h\Lambda_h(t)F_h(t), \qquad t \in [0,T], \tag{3.2}$$

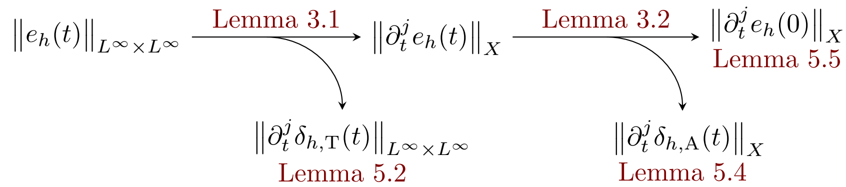


FIGURE 1. Strategy of the proof of Theorem 2.8.

with initial value $y_h(0) = y_h^0$. Thus, we conclude that the discrete error e_h solves the evolution equation

$$\mathrm{T}_h \Lambda_h(t) \partial_t e_h(t) = e_h(t) + \delta_{h,\mathrm{T}}(t), \qquad t \in [0,T], \tag{3.3}$$

with initial value $e_h(0) = e_h^0 = J_h y_0 - y_h(0)$ and the defect

$$\delta_{h,\mathrm{T}}(t) = \big(\mathrm{T}_h J_h - J_h \mathrm{T}\big)\Lambda(t)\partial_t y(t) + J_h \mathrm{T}\Lambda(t)F(t) - \mathrm{T}_h \Lambda_h(t) F_h(t). \tag{3.4}$$

As illustrated in Figure 1, the proof of Theorem 2.8 mainly consists of two steps. First, in Lemma 3.1 we exchange the maximum norm of $e_h(t)$ for bounds of time derivatives of $e_h(t)$ in X_h. To do so, we use (3.3) and Lemma 2.7, i.e., we rely on the property of the solution operator to gain integrability as sketched in (2.26).

Next, in Lemma 3.2 we trace back the time derivatives of $e_h(t)$ to time derivatives of the initial error $e_h(0)$, which can be bounded due to the specific choice (2.24) of the discrete initial value y_h^0. Here, we obtain from (2.2) and (2.23) for the discrete error e_h the evolution equation

$$\Lambda_h(t)\partial_t e_h(t) = \mathrm{A}_h e_h(t) + \delta_{h,\mathrm{A}}(t), \qquad t \in [0,T], \tag{3.5}$$

with the defect

$$\delta_{h,\mathrm{A}}(t) = \big(J_h \mathrm{A} - \mathrm{A}_h J_h\big) y(t) + J_h \Lambda(t) F(t) - \Lambda_h(t) F_h(t). \tag{3.6}$$

Note that we have the relation $\delta_{h,\mathrm{A}} = \mathrm{A}_h \delta_{h,\mathrm{T}}$. Moreover, we emphasize that this defect was already studied in the unified error analysis provided in [18]. However, here we also have to bound time derivatives of $\delta_{h,\mathrm{T}}$ and $\delta_{h,\mathrm{A}}$. We postpone the derivation of these bounds as well as the estimates for the time derivatives of the initial error to Section 5.

3.2. **Proof of Theorem 2.8.** Our first result shows how to bound the error in the maximum norm in terms of time derivatives of the error in the energy norm.

Lemma 3.1. *Let the assumptions of Theorem 2.8 hold. Then, it holds*

$$\max_{[0,T]} \|e_h\|_{L^\infty \times L^\infty} \lesssim \sum_{j=0}^{2} \max_{[0,T]} \left\|\partial_t^j \delta_{h,\mathrm{T}}\right\|_{L^\infty \times L^\infty} + \sum_{j=0}^{2} \left\|\partial_t^{j+1} e_h\right\|_{L^\infty(X)}.$$

Proof. From (3.3) we obtain for $j = 0,1,2$ the relation

$$\partial_t^j e_h(t) = -\partial_t^j \delta_{h,\mathrm{T}}(t) + \sum_{\ell=0}^{j} \binom{j}{\ell} \mathrm{T}_h \partial_t^{j-\ell} \Lambda_h(t) \partial_t^{\ell+1} e_h(t)\,. \tag{3.7}$$

We choose $(p_0, p_1, p_2, p_3, p_4) = (\infty, \infty, 4, 4, 2)$ and, since $0 < \frac{1}{p_{n+2}} - \frac{1}{p_n} < \frac{1}{N}$, taking norms gives by Lemma 2.7 the estimate

$$\begin{aligned}\left\|\partial_t^j e_h(t)\right\|_{L^{p_n}\times L^{p_{n+1}}} &\lesssim \left\|\partial_t^j \delta_{h,\mathrm{T}}(t)\right\|_{L^{p_n}\times L^{p_{n+1}}} + \sum_{\ell=0}^{j} \left\|\mathrm{T}_h \partial_t^{j-\ell} \Lambda_h(t) \partial_t^{\ell+1} e_h(t)\right\|_{L^{p_n}\times L^{p_{n+1}}} \\ &\lesssim \left\|\partial_t^j \delta_{h,\mathrm{T}}(t)\right\|_{L^\infty\times L^\infty} + \sum_{\ell=0}^{j} \left\|\partial_t^{\ell+1} e_h(t)\right\|_{L^{p_{n+1}}\times L^{p_{n+2}}}.\end{aligned}$$

Resolving this, we arrive at

$$\|e_h(t)\|_{L^\infty\times L^\infty} \lesssim \sum_{j=0}^{2} \left\|\partial_t^j \delta_{h,\mathrm{T}}(t)\right\|_{L^\infty\times L^\infty} + \sum_{j=0}^{2} \left\|\partial_t^{j+1} e_h(t)\right\|_{L^4\times L^2}.$$

Finally, we use Sobolev's embedding to obtain

$$\|\xi_h\|_{L^4\times L^2} \lesssim \|\xi_h\|_X, \quad \xi_h \in X_h,$$

which concludes the proof. □

In the following lemma, we provide the bounds of the time derivatives appearing in Lemma 3.1 in the X norm using the initial errors and certain defects.

Lemma 3.2. *Let the assumptions of Theorem 2.8 hold. Then, there is a constant $C > 0$ independent of h such that for $1 \le j \le 3$ we have*

$$\left\|\partial_t^j e_h\right\|_{L^\infty(X)}^2 \lesssim \mathrm{e}^{CT} \sum_{\ell=1}^{j} \Big(\left\|\partial_t^\ell e_h(0)\right\|_X^2 + \left\|\partial_t^\ell \delta_{h,\mathrm{A}}\right\|_{L^\infty(X)}^2 \Big).$$

Proof. In the following, we prove for $1 \le j \le 3$ the estimate

(3.8)
$$\left\|\partial_t^j e_h\right\|_{L^\infty(X)}^2 \lesssim (1+T)\mathrm{e}^{CT} \left(\left\|\partial_t^j e_h(0)\right\|_X^2 + \left\|\partial_t^j \delta_{h,\mathrm{A}}\right\|_{L^\infty(X)}^2 + \sum_{\ell=1}^{j-1} \left\|\partial_t^\ell e_h\right\|_{L^\infty(X)}^2 \right).$$

The result then follows from using this bound recursively. In the following, we often suppress the time arguments to increase the readability.

To prove (3.8), we first obtain by taking the derivative of (3.5) with respect to time

$$\sum_{\ell=0}^{j} \binom{j}{\ell} \partial_t^{j-\ell} \Lambda_h \partial_t^{\ell+1} e_h - \mathrm{A}_h \partial_t^j e_h = \partial_t^j \delta_{h,\mathrm{A}},$$

for $j = 0, \dots, 3$. Taking the inner product with $\partial_t^j e_h$ gives

$$\begin{aligned}\left(\Lambda_h \partial_t^{j+1} e_h \mid \partial_t^j e_h\right)_X &= \left(\mathrm{A}_h \partial_t^j e_h \mid \partial_t^j e_h\right)_X + \left(\partial_t^j \delta_{h,\mathrm{A}} \mid \partial_t^j e_h\right)_X \\ &\quad - \sum_{\ell=0}^{j-1} \binom{j}{\ell} \left(\partial_t^{j-\ell} \Lambda_h \partial_t^{\ell+1} e_h \mid \partial_t^j e_h\right)_X.\end{aligned}$$

Since A_h is skew-symmetric with respect to the inner product of X_h, we obtain with the triangle inequality and Young' inequality the bound

$$2\left(\Lambda_h \partial_t^{j+1} e_h \mid \partial_t^j e_h\right)_X \le \left\|\partial_t^j \delta_{h,\mathrm{A}}\right\|_X^2 + 2^j \left\|\partial_t^j e_h\right\|_X^2 + \sum_{\ell=0}^{j-1} \binom{j}{\ell} \left\|\partial_t^{j-\ell} \Lambda_h \partial_t^{\ell+1} e_h\right\|_X^2.$$

In particular, due to the boundedness of Λ_h by Lemma 2.6 and the corresponding time derivatives, we conclude

$$2\left(\Lambda_h \partial_t^{j+1} e_h \mid \partial_t^j e_h\right)_X \leq \left\|\partial_t^j \delta_{h,\mathrm{A}}\right\|_X^2 + C_j \left\|\partial_t^j e_h\right\|_X^2 + \widehat{C}_j \sum_{\ell=1}^{j-1} \left\|\partial_t^\ell e_h\right\|_X^2, \tag{3.9}$$

with the constants

$$C_j = 2^j + j \left\|\partial_t \Lambda_h\right\|_{\mathcal{L}(X_h)}^2, \qquad \widehat{C}_j = \max_{\ell=0,\dots,j-1} \binom{j}{\ell} \left\|\partial_t^{j-\ell} \Lambda_h\right\|_{L^\infty(\mathcal{L}(X_h))}^2.$$

Note that these constants are bounded independently of $j \leq 3$ by C_3 and $\widehat{C}_3$, respectively.

We rely on (3.9) to bound the first term on the right-hand side of

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\|\partial_t^j e_h\right\|_{\Lambda_h(t)}^2 = 2\left(\Lambda_h \partial_t^{j+1} e_h \mid \partial_t^j e_h\right)_X + \left(\partial_t \Lambda_h \partial_t^j e_h \mid \partial_t^j e_h\right)_X.$$

Moreover, integration in time, using the boundedness of $\partial_t \Lambda_h$ for the second term, and the norm equivalence (2.21) yields

$$\begin{aligned}\left\|\partial_t^j e_h\right\|_{\Lambda_h(t)}^2 \lesssim \left\|\partial_t^j e_h(0)\right\|_{\Lambda_h(0)}^2 + t\left\|\partial_t^j \delta_{h,\mathrm{A}}\right\|_{L^\infty(X)}^2 + t\sum_{\ell=1}^{j-1} \left\|\partial_t^\ell e_h\right\|_{L^\infty(X)}^2 \\ + \int_0^t \left\|\partial_t^j e_h(s)\right\|_{\Lambda_h(s)}^2 \mathrm{d}s.\end{aligned}$$

Finally, the Gronwall inequality implies for all $t \in [0,T]$

$$\left\|\partial_t^j e_h(t)\right\|_{\Lambda_h(t)}^2 \lesssim \mathrm{e}^{Ct}\left(\left\|\partial_t^j e_h(0)\right\|_{\Lambda_h(0)}^2 + t\left\|\partial_t^j \delta_{h,\mathrm{A}}\right\|_{L^\infty(X)}^2 + \sum_{\ell=1}^{j-1} \left\|\partial_t^\ell e_h\right\|_{L^\infty(X)}^2\right)$$

and (3.8) follows with (2.21). □

With these preparations we can prove our first main result.

Proof of Theorem 2.8. Using the decomposition (3.1) and the stability of the lift in (2.7), we estimate with the approximation properties derived in (2.9) and (2.12)

$$\begin{aligned}\left\|y(t) - \boldsymbol{\mathcal{L}_h} y_h(t)\right\|_{L^\infty \times L^\infty} &\leq \left\|e_{J_h}(t)\right\|_{L^\infty \times L^\infty} + C_{\boldsymbol{\mathcal{L}_h},p} \left\|e_h(t)\right\|_{L^\infty \times L^\infty} \\ &\leq Ch^k \left\|y(t)\right\|_{W^{k+1,\infty} \times W^{k,\infty}} + C_{\boldsymbol{\mathcal{L}_h},p} \left\|e_h(t)\right\|_{L^\infty \times L^\infty},\end{aligned}$$

and apply Lemmas 3.1 and 3.2. The remaining defects and errors in the initial values are bounded in Lemmas 5.2, 5.4 and 5.5. □

4. Analysis of the full discretization

In this section, we establish the proof of Theorem 2.9. The strategy is very similar to the one in Section 3, see Figure 2, where we replace the continuous by discrete derivatives. Hence, after introducing some useful calculus, we explain the adapted strategy.

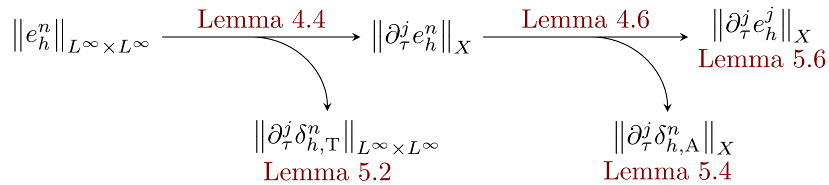


FIGURE 2. Strategy of the proof of Theorem 2.9.

4.1. **Calculus for discrete derivatives.** We first need some auxiliary results for the discrete derivatives defined in (2.28). A straightforward calculation yields the following.

Lemma 4.1. *It holds the discrete product rule*

$$\partial_\tau\big(\varphi^n\psi^n\big) = (\partial_\tau\varphi^n)\psi^n + \varphi^{n-1}(\partial_\tau\psi^n)$$

and also the more general discrete Leibniz rule

$$\partial_\tau^j\big(\varphi^n\psi^n\big) = \sum_{\ell=0}^{j}\binom{j}{\ell}\big(\partial_\tau^{j-\ell}\varphi^{n-\ell}\big)\big(\partial_\tau^\ell\psi^n\big), \quad j \ge 0.$$

In order to mimic the strategy of the proof of Theorem 2.8, we state the well-known discrete version of the fundamental theorem of calculus and a direct consequence of a discrete Gronwall lemma.

Lemma 4.2. *Let (φ^n) be a sequence in a Hilbert space with inner product $(\cdot \mid \cdot)$, and let $k_0 \in \mathbb{N}$.*

(a) *For any $M \ge k_0$, it holds*

$$\frac{1}{2}\left\|\varphi^M\right\|^2 \le \frac{1}{2}\left\|\varphi^{k_0-1}\right\|^2 + \tau\sum_{j=k_0}^{M}\left(\partial_\tau\varphi^j \mid \varphi^j\right).$$

(b) *If there are constants $\alpha, \beta_1, \beta_2 \ge 0$ such that*

$$\left(\partial_\tau\varphi^j \mid \varphi^j\right) \le \alpha^2 + \beta_1\left\|\varphi^{j-1}\right\|^2 + \beta_2\left\|\varphi^j\right\|^2, \quad j \ge k_0, \tag{4.1}$$

holds, then for $\tau \le \frac{1}{4(\beta_1+\beta_2)}$ and $M \ge k_0$ we have

$$\left\|\varphi^M\right\| \le \Big(\sqrt{1+2\tau\beta_1}\left\|\varphi^{k_0-1}\right\| + \sqrt{2t_N}\alpha\Big)\mathrm{e}^{2(\beta_1+\beta_2)t_N}.$$

Proof. Part (a) is for example shown in [20, Lemma 4.2]. Inserting (4.1) in (a) yields

$$\begin{aligned}\left\|\varphi^N\right\|^2 &\le \left\|\varphi^{k_0-1}\right\|^2 + 2\tau\sum_{j=k_0}^{N}\Big(\alpha^2 + \beta_1\left\|\varphi^{j-1}\right\|^2 + \beta_2\left\|\varphi^j\right\|^2\Big)\\ &\le \big(1+2\tau\beta_1\big)\left\|\varphi^{k_0-1}\right\|^2 + 2t_N\alpha^2 + 2(\beta_1+\beta_2)\tau\sum_{j=k_0}^{N}\left\|\varphi^j\right\|^2\end{aligned}$$

and by a Gronwall argument, see, e.g., [26, Lemma 1], we obtain

$$\left\|\varphi^N\right\|^2 \le \Big(\big(1+2\tau\beta_1\big)\left\|\varphi^{k_0-1}\right\|^2 + 2t_N\alpha^2\Big)\mathrm{e}^{4(\beta_1+\beta_2)t_N}.$$

Taking roots yields the assertion. □

We conclude with a useful bound which relates the discrete derivatives to their continuous limit.

Lemma 4.3. *Let Z be some Banach space, $j \geq 1$ and $x\colon [0,T] \to Z$ be j-times differentiable with bounded derivatives, then*

$$\left\|\partial_\tau^j x(t^n)\right\|_Z \leq \sup_{t\in[t^{n-j},t^n]} \left\|\partial_t^j x(t)\right\|_Z .$$

Proof. This simply follows from an iterative application of the fundamental theorem of calculus. □

4.2. **Proof of Theorem 2.9.** As in (3.1), we are interested in bounds on the discrete error

$$e_h^n = J_h y(t^n) - y_h^n,$$

and derive for the exact solution inserted in the numerical scheme similar to (3.2)

$$\mathrm{T}_h \Lambda_h^n J_h \partial_\tau y(t^n) = J_h y(t^n) + \mathrm{T}_h \Lambda_h^n F_h^n + \delta_{h,\mathrm{T}}^n$$

with a defect of the form, using (3.4),

$$\delta_{h,\mathrm{T}}^n = \delta_{h,\mathrm{T}}(t^n) + \mathrm{T}_h J_h \Lambda^n \big(\partial_\tau y(t^n) - \partial_t y(t^n)\big). \tag{4.2}$$

From this we obtain the fully discrete error equation

$$\mathrm{T}_h \Lambda_h(t^n) \partial_\tau e_h^n = e_h^n + \delta_{h,\mathrm{T}}^n. \tag{4.3}$$

For the estimates in the energy we need the equivalent formulation involving the operator A_h. To this end, we insert $J_h y$ into (2.27) and obtain

$$\Lambda_h(t^n)\partial_\tau y(t^n) = \mathrm{A}_h y(t^n) + F_h^n + \delta_{h,\mathrm{A}}^n$$

with the defect, using (3.6),

$$\delta_{h,\mathrm{A}}^n = \delta_{h,\mathrm{A}}(t^n) + J_h \Lambda^n \big(\partial_\tau y(t^n) - \partial_t y(t^n)\big). \tag{4.4}$$

This gives us the second version of the error recursion

$$\Lambda_h(t^n)\partial_\tau e_h^n = \mathrm{A}_h e_h^n + \delta_{h,\mathrm{A}}^n. \tag{4.5}$$

Starting from (4.3), we obtain the following bound as a discrete counterpart to Lemma 3.1.

Lemma 4.4. *Let the assumptions of Theorem 2.9 hold. Then, there exists a constant $C > 0$ independent of h, τ, and n such that*

$$\|e_h^n\|_{L^\infty\times L^\infty} \leq C \sum_{j=0}^{2} \left\|\partial_\tau^j \delta_{h,\mathrm{T}}^n\right\|_{L^\infty\times L^\infty} + C\sum_{j=0}^{2} \left\|\partial_\tau^{j+1} e_h^n\right\|_X$$

holds for $n \geq 3$.

Remark 4.5. From this lemma it becomes clear that this technique does not provide bounds for $n = 0, 1, 2$, since we can evaluate $\partial_\tau^j e_h^n$ only for $n \geq j$.

Proof. As in (3.7), we obtain from (4.3) and Lemma 4.1 for $j = 0, 1, 2$

$$\begin{aligned}\partial_\tau^j e_h^n &= \partial_\tau^j\Big(-\delta_{h,\mathrm{T}}^n + \mathrm{T}_h \Lambda_h(t^n)\partial_\tau e_h^n\Big)\\ &= -\partial_\tau^j \delta_{h,\mathrm{T}}^n + \mathrm{T}_h \sum_{\ell=0}^{j} \binom{j}{\ell} \partial_\tau^{j-\ell} \Lambda_h(t^{n-\ell}) \partial_\tau^{\ell+1} e_h^n\end{aligned}$$

and the proof follows along the lines of Lemma 3.1. ☐

The next step is to establish the discrete analogue to Lemma 3.2 where the discrete derivatives are bounded in terms of discrete derivatives of the initial error.

Lemma 4.6. *Let the assumptions of Theorem 2.9 hold. Then, we have*

$$\left\|\partial_\tau^j e_h^n\right\|_X^2 \leq C(1+T)\mathrm{e}^{CT} \sum_{\ell=1}^{j} \Big(\left\|\partial_\tau^\ell e_h^\ell\right\|_X^2 + \left\|\partial_\tau^\ell \delta_{h,\mathrm{A}}^n\right\|_{\ell^\infty(X)}^2 \Big),$$

for $1 \leq j \leq 3$ *and* $n \geq j+1$.

Proof. As in the proof of Lemma 3.2, we provide the bound

$$\left\|\partial_\tau^j e_h^n\right\|_X^2 \leq C(1+T)\mathrm{e}^{CT} \Bigg(\left\|\partial_\tau^j e_h^j\right\|_X^2 + \left\|\partial_\tau^j \delta_{h,\mathrm{A}}^n\right\|_{\ell^\infty(X)}^2 + \sum_{\ell=1}^{j-1} \left\|\partial_\tau^\ell e_h^n\right\|_X^2 \Bigg), \tag{4.6}$$

cf. (3.8). Using this estimate recursively directly yields the assertion.

To do so, we apply the bounds from Lemma 4.2 for $\varphi^n = \Lambda_h^{1/2}(t^n)\partial_\tau^j e_h^n$. In particular, we first study the term

$$(\partial_\tau \varphi^n \mid \varphi^n)_X = \left(\Lambda_h^n \partial_\tau^{j+1} e_h^n \mid \partial_\tau^j e_h^n\right)_X + \Big(\big(\partial_\tau \Lambda_h^{1/2}(t^n)\big) \partial_\tau^j e_h^{n-1} \mid \varphi^n \Big)_X ,$$

where we used Lemma 4.1 and the fact that $\Lambda_h^{1/2}$ is self-adjoint in X_h. Due to Assumption 2.1 and Young's inequality, this implies

$$(\partial_\tau \varphi^n \mid \varphi^n)_X \leq \left(\Lambda_h^n \partial_\tau^{j+1} e_h^n \mid \partial_\tau^j e_h^n\right)_X + C\left\|\varphi^n\right\|_X^2 + C\left\|\varphi^{n-1}\right\|_X^2 . \tag{4.7}$$

For the first term, we obtain with Lemma 4.1 applied to (4.5)

$$\Lambda_h^n \partial_\tau^{j+1} e_h^n = \mathrm{A}_h \partial_\tau^j e_h^n + \partial_\tau^j \delta_{h,\mathrm{A}}^n - \sum_{\ell=0}^{j-1} \binom{j}{\ell} \big(\partial_\tau^{j-\ell} \Lambda_h^{n-\ell}\big)\big(\partial_\tau^{\ell+1} e_h^n\big).$$

Thus, taking the inner product in X_h with $\partial_\tau^j e_h^n$ yields as in (3.9) the estimate

$$2\left(\Lambda_h^n \partial_\tau^{j+1} e_h^n \mid \partial_\tau^j e_h^n\right)_{X_h} \leq \left\|\partial_\tau^j \delta_{h,\mathrm{A}}^n\right\|_X^2 + C\left\|\varphi^n\right\|_X^2 + C \sum_{\ell=1}^{j-1} \left\|\partial_\tau^\ell e_h^n\right\|_X^2 .$$

Using this bound in (4.7) together with the discrete Gronwall lemma in Lemma 4.2, we obtain (4.6). ☐

Hence, we conclude our second main result.

Proof of Theorem 2.9. Using a decomposition analogous to (3.1), we estimate as in the proof of Theorem 2.8

$$\begin{aligned} \|y(t^n) - \boldsymbol{\mathcal{L}_h} y_h^n\|_{L^\infty \times L^\infty} &\leq \|e_{J_h}(t^n)\|_{L^\infty \times L^\infty} + C_{\boldsymbol{\mathcal{L}_h},p} \|e_h^n\|_{L^\infty \times L^\infty} \\ &\leq Ch^k \|y(t^n)\|_{W^{k+1,\infty} \times W^{k,\infty}} + C_{\boldsymbol{\mathcal{L}_h},p} \|e_h^n\|_{L^\infty \times L^\infty} , \end{aligned}$$

and apply Lemmas 4.4 and 4.6 for $n \geq 3$. The remaining defects and errors in the initial values are bounded in Lemmas 5.2, 5.4 and 5.6. ☐

5. Bounds on the defects and initial conditions

In this section, we provide all bounds missing in the proofs of Sections 3 and 4. Throughout, we mostly omit the time dependency for the sake of readability. Further, we take the assumptions of Section 2 as given and will only be precise about the regularity of the solution y and the right-hand side f.

5.1. **Estimates of the defects.**

We first provide certain approximation properties in the maximum norm which are used for the defects.

Lemma 5.1. *Let $\xi \in W^{k+1,\infty}(\Omega) \times W^{k,N+1}(\Omega)$ and $f \in C^2([0,T], W^{k+1,\infty}(\Omega))$. Then, the discrete operators introduced in Section 2 satisfy for $j \in \{0,1,2\}$ the bounds*

$$\big\| \big(\mathrm{T}_h J_h - J_h \mathrm{T}\big)\xi \big\|_{L^\infty \times L^\infty} \lesssim h^k \, \|\xi\|_{W^{k+1,\infty}\times W^{k,N+1}} \,, \tag{5.1}$$

$$\big\| \partial_t^j F_h - J_h \partial_t^j F \big\|_{L^\infty \times L^\infty} \lesssim h^k \big\| \partial_t^j f \big\|_{W^{k+1,\infty}} \,. \tag{5.2}$$

Moreover, for $\xi \in \big(H^4(\Omega) \cap \mathcal{D}(\mathrm{L})\big) \times \big(W^{2,\infty}(\Omega) \cap \mathcal{D}(\mathrm{L})\big)$ and the inhomogeneity. $f \in C\big([0,T], H^4(\Omega) \cap \mathcal{D}(\mathrm{L})\big)$ we have the estimates

$$\big\| \mathrm{A}_h^2 F_h - J_h \mathrm{A}^2 F \big\|_{L^\infty \times L^\infty} \lesssim \|f\|_{H^4} \,, \tag{5.3}$$

$$\big\| \big(\mathrm{A}_h J_h - J_h \mathrm{A}\big)\xi \big\|_{L^\infty \times L^\infty} \lesssim \|\xi\|_{H^4 \times W^{1,\infty}} \,. \tag{5.4}$$

Proof. We have for $\xi = (\varphi, \psi) \in W^{k+1,\infty}(\Omega) \times W^{k,N+1}(\Omega)$

$$\big(\mathrm{T}_h J_h - J_h \mathrm{T}\big)\xi = \begin{pmatrix} \big(S_h I_h - \mathcal{L}_h^{V*} S\big)\psi \\ \big(\mathcal{L}_h^{V*} - I_h\big)\varphi \end{pmatrix} , \tag{5.5}$$

and the second component is bounded by (2.14). Similar to (A.2), we employ the inverse estimate (2.15) and compute with (2.6) and (2.22)

$$\begin{aligned} \big\| \big(S_h I_h - \mathcal{L}_h^{V*} S\big)\psi \big\|_{L^\infty} &\lesssim h^{-N/6} \sup_{\|\phi_h\|_{V_h}=1} \big(\big(S_h I_h - \mathcal{L}_h^{V*} S\big)\psi \mid \phi_h \big)_{V_h} \\ &= h^{-N/6} \sup_{\|\phi_h\|_{V_h}=1} \big(\big(I_h - \mathcal{L}_h^{H*}\big)\psi \mid \phi_h \big)_{H_h} \\ &\leq h^k \, \|\psi\|_{H^{k+1}} \end{aligned}$$

where we used the approximation property in (2.11) and hence, (5.1) follows. Similarly, (5.2) follows for $j = 0, 1, 2$ from

$$\partial_t^j F_h - J_h \partial_t^j F = \begin{pmatrix} 0 \\ (\mathcal{L}_h^{V*} - I_h)\partial_t^j f \end{pmatrix} \tag{5.6}$$

by (2.14). To prove (5.3), we first observe

$$\mathrm{A}_h^2 F_h - J_h \mathrm{A}^2 F = \begin{pmatrix} 0 \\ (\mathrm{L}_h \mathcal{L}_h^{V*} - I_h \mathrm{L}) f \end{pmatrix} . \tag{5.7}$$

Thus, the inverse estimate (2.15) yields

$$\begin{aligned} \big\| \mathrm{A}_h^2 F_h - J_h \mathrm{A}^2 F \big\|_{L^\infty \times L^\infty} &\lesssim h^{-N/2} \, \big\| (\mathrm{L}_h \mathcal{L}_h^{V*} - I_h \mathrm{L}) f \big\|_{L^2} \\ &\lesssim h^{-N/2} \, \big\| (\mathcal{L}_h^{H*} - I_h) \mathrm{L} f \big\|_{L^2} \\ &\lesssim h^{1/2} \, \|\mathrm{L} f\|_{H^2} \lesssim \|f\|_{H^4} \,. \end{aligned}$$

Note that we used the identity $\mathrm{L}_h\mathcal{L}_h^{V*} = \mathcal{L}_h^{H*}\mathrm{L}$ to derive the second estimate and (2.11) for the third. Due to

$$(\mathrm{A}_h J_h - J_h\mathrm{A})\xi = \begin{pmatrix} (I_h - \mathcal{L}_h^{V*})\psi \\ (\mathrm{L}_h\mathcal{L}_h^{V*} - I_h\mathrm{L})\varphi \end{pmatrix} ,$$

the last estimate (5.4) follows with similar arguments. □

From these approximation properties we conclude the bounds on the defects appearing in Lemmas 3.1 and 4.4.

Lemma 5.2. *Let the solution $y \in C^3([0,T], W^{k+1,\infty}(\Omega) \times W^{k+1,N+1}(\Omega))$ and the right-hand side $f \in C^2([0,T], W^{k+1,\infty}(\Omega))$.*

(a) *The defect $\delta_{h,\mathrm{T}}(t)$ introduced in (3.4) satisfies for $j \in \{0,1,2\}$ and $t \in [0,T]$*

$$\big\|\partial_t^j \delta_{h,\mathrm{T}}(t)\big\|_{L^\infty\times L^\infty} \lesssim h^k .$$

(b) *If in addition $y \in C^4([0,T], W^{1,\infty}(\Omega) \times W^{1,N+1}(\Omega))$, then the defect $\delta_{h,\mathrm{T}}^n$ introduced in (4.2) satisfies for $j \in \{0,1,2\}$ and $n \geq j$*

$$\big\|\partial_\tau^j \delta_{h,\mathrm{T}}^n\big\|_{L^\infty\times L^\infty} \lesssim \tau + h^k .$$

Proof. (a) We decompose the defect $\delta_{h,\mathrm{T}} = \delta_{h,\mathrm{T}}^1 + \delta_{h,\mathrm{T}}^2$ introduced in (3.4) with

$$\delta_{h,\mathrm{T}}^1 = \big(\mathrm{T}_h J_h - J_h\mathrm{T}\big)\Lambda\partial_t y\,, \qquad \delta_{h,\mathrm{T}}^2 = J_h\mathrm{T}\Lambda(t)F(t) - \mathrm{T}_h\Lambda_h(t)F_h(t)\,,$$

and first consider $\delta_{h,\mathrm{T}}^1$. Differentiating gives for $j \in \{0,1,2\}$

$$\partial_t^j \delta_{h,\mathrm{T}}^1 = \big(\mathrm{T}_h J_h - J_h\mathrm{T}\big)\sum_{\ell=0}^{j}\binom{j}{\ell}\partial_t^{j-\ell}\Lambda\,\partial_t^{\ell+1} y\,,$$

so that (5.1) together with Lemma 2.3 implies

$$\big\|\partial_t^j \delta_{h,\mathrm{T}}^1\big\|_{L^\infty\times L^\infty} \lesssim h^k \sum_{\ell=0}^{j}\big\|\partial_t^{\ell+1} y\big\|_{W^{k+1,\infty}\times W^{k+1,N+1}} .$$

With similar arguments, we obtain for $j \in \{0,1,2\}$

$$\partial_t^j \delta_{h,\mathrm{T}}^2 = \big(J_h\mathrm{T} - \mathrm{T}_h J_h\big)\sum_{\ell=0}^{j}\binom{j}{\ell}\partial_t^{j-\ell}\Lambda\,\partial_t^{\ell} F + \mathrm{T}_h\sum_{\ell=0}^{j}\binom{j}{\ell}\partial_t^{j-\ell}\Lambda_h\big(J_h\partial_t^\ell F - \partial_t^\ell F_h\big)\,.$$

Thus, (5.1) and (5.2) together with Lemma 2.6 imply

$$\Big\|\partial_t^j \delta_{h,\mathrm{T}}^2\Big\|_{L^\infty\times L^\infty} \lesssim h^k \sum_{\ell=0}^{j}\big\|\partial_t^\ell f\big\|_{W^{k+1,\infty}} .$$

(b) Thanks to Lemma 4.3, the estimate from part (a) extends to this case and we only have to provide the following bound for the additional defect

$$\big\|\mathrm{T}_h J_h\partial_\tau^j\Big(\Lambda^n\big(\partial_\tau y(t^n) - \partial_t y(t^n)\big)\Big)\big\|_{L^\infty\times L^\infty} \lesssim \tau\big\|\partial_t^{j+2} y\big\|_{W^{1,\infty}\times W^{1,\infty}}. \tag{5.8}$$

We first note that it holds

$$\partial_\tau y(t^n) - \partial_t y(t^n) = \tau\int_0^1 (-s)\partial_t^2 y(t^{n-1} + \tau s)\,\mathrm{d}s\,, \tag{5.9}$$

and hence, we obtain

$$\begin{aligned}&\partial_\tau^j\Big(\Lambda^n\big(\partial_\tau y(t^n)-\partial_t y(t^n)\big)\Big)\\ &=\tau\sum_{\ell=0}^{j}\binom{j}{\ell}\partial_\tau^{j-\ell}\Lambda^{n-\ell}\int_0^1(-s)\partial_\tau^{\ell}\partial_t^2 y(t^{n-1}+\tau s)\,\mathrm{d}s\,.\end{aligned} \tag{5.10}$$

The statement in (5.8) then follows from Assumption 2.1, (2.16), Lemma 2.7, and Lemma 4.3. □

In the next step, we consider the time derivatives of the error which we estimate in the energy norm. We again provide some approximation properties and recall

$$k^*=\max\{k,2\}.$$

Additionally, we define the operator

$$\widetilde{\Lambda^n}:=\begin{pmatrix}\lambda(t^n) & 0\\ 0 & \mathrm{Id}\end{pmatrix} \tag{5.11}$$

and similarly, $\widetilde{\Lambda_h^n}$, $(\widetilde{\Lambda_h^n})^{-1}$, and $(\widetilde{\Lambda^n})^{-1}$.

Lemma 5.3. *Let $\xi\in H^{k+1}(\Omega)\times H^{k^*}(\Omega)$ and $f\in C^2([0,T],H^{k^*}(\Omega))$. Then, the discrete operators introduced in Section 2 satisfy for $j\in\{0,1,2\}$ the bounds*

$$\big\|\big(\widetilde{\Lambda}J_h-J_h\widetilde{\Lambda_h}\big)\xi\big\|_X\lesssim h^k\,\|\xi\|_{H^{k+1}\times H^{k^*}}\ , \tag{5.12}$$

$$\big\|\big(\mathrm{T}_hJ_h-J_h\mathrm{T}\big)\xi\big\|_X\lesssim h^k\,\|\xi\|_{H^{k+1}\times H^{k^*}}\ , \tag{5.13}$$

$$\big\|\partial_t^jF_h-J_h\partial_t^jF\big\|_X\lesssim h^k\big\|\partial_t^jf\big\|_{H^{k^*}}\ . \tag{5.14}$$

If $\xi\in H^{k+1}(\Omega)\times H^{k^*}(\Omega)\cap\mathcal{D}(\mathrm{A})$ and $f\in C([0,T],H^{k^*+2}(\Omega))\cap\mathcal{D}(\Delta)$, then

$$\big\|\mathrm{A}_h^2F_h-J_h\mathrm{A}^2F\big\|_X\lesssim h^k\,\|f\|_{H^{k^*+2}}\ , \tag{5.15}$$

$$\big\|\big(\mathrm{A}_hJ_h-J_h\mathrm{A}\big)\xi\big\|_X\lesssim h^k\,\|\xi\|_{H^{k^*+2}\times H^{k+1}}\ . \tag{5.16}$$

Proof. Let again $\xi=(\varphi,\psi)$. For the first term, we have

$$\big\|\big(\widetilde{\Lambda}J_h-J_h\widetilde{\Lambda_h}\big)\xi\big\|_X=\big\|\lambda_h\mathcal{L}_h^{V*}\varphi-\mathcal{L}_h^{V*}(\lambda\varphi)\big\|_{V_h}=\big\|I_h(\lambda\,\mathcal{L}_h\mathcal{L}_h^{V*}\varphi)-\mathcal{L}_h^{V*}(\lambda\varphi)\big\|_{V_h}$$

and basic manipulations together with (2.19) and (2.14) yield (5.12). For the other estimates, we proceed analogously to the proof of Lemma 5.1 using the representation derived there. From (5.5) we derive by the stability of S_h in V_h together with (2.9) and (2.12) the bound (5.13). By the same arguments, (5.6) implies (5.14).

Starting from (5.7), we follow the lines of Lemma 5.1 without the inverse estimate to derive (5.15). Finally, from [18, Lem. 4.7] we obtain the bound

$$\big\|\big(\mathrm{A}_hJ_h-J_h\mathrm{A}\big)\xi\big\|_X\lesssim\|(\mathrm{Id}-I_h)\varphi\|_{H^1}+\|(\mathrm{Id}-I_h)\psi\|_{H^1}+\|(\mathrm{Id}-I_h)\mathrm{L}\varphi\|_{L^2}\,,$$

and (5.16) follows from the bounds in (2.9). □

With these approximations at hand, we can finally provide bounds for the defects from Lemmas 3.2 and 4.6.

Lemma 5.4. *Let the solution $y\in C^3\big([0,T],\big(H^{k^*+2}(\Omega)\cap\mathcal{D}(\mathrm{L})\big)\times\big(H^{k+1}(\Omega)\cap\mathcal{D}(\mathrm{L})\big)\big)$ and the right-hand side $f\in C^3([0,T],H^{k^*}(\Omega))$.*

(a) *The defect defined in* (3.6) *satisfies for* $j \in \{0,1,2,3\}$ *and* $t \in [0,T]$

$$\left\| \partial_t^j \delta_{h,\mathrm{A}}(t) \right\|_X \lesssim h^k .$$

(b) *If in addition* $y \in C^5([0,T], H^1(\Omega) \times H^2(\Omega))$*, then the defect defined in* (4.4) *satisfies for* $j \in \{0,1,2,3\}$ *and* $n \geq j$

$$\left\| \partial_\tau^j \delta_{h,\mathrm{A}}^n \right\|_X \lesssim \tau + h^k .$$

Proof. (a) We estimate the defect with the help of Lemma 5.3. We write with (2.20)

$$J_h \Lambda(t) F(t) - \Lambda_h(t) F_h(t) = \Lambda_h(t)\big(J_h F(t) - F_h(t)\big),$$

and employ Lemma 2.6 and the bounds (5.14) and (5.16) to obtain

$$\left\| \partial_t^j \delta_{h,\mathrm{A}}(t) \right\|_X \leq C h^k \big(\left\| \partial_t^j y \right\|_{H^{k^*+2} \times H^{k+1}} + \left\| \partial_t^j f \right\|_{H^{k^*}} \big).$$

(b) Thanks to Lemma 4.3, it remains to show

$$\left\| \partial_\tau^j J_h \Lambda^n \big(y(t^n) - \partial_t y(t^n) \big) \right\|_X \lesssim \tau \left\| \partial_t^{j+2} y \right\|_{H^1 \times H^2} .$$

Using the continuity of J_h and following the lines of (5.10), immediately yields the claim. □

5.2. **Errors of the differentiated initial values.** The last part to prove Theorem 2.8 and Theorem 2.9 is to bound the initial error in Lemma 3.2 and Lemma 4.4. We recall the preconditioned initial values defined in (2.24)

$$y_h(0) = y_h^0 = \mathrm{T}_h \Lambda_h(0) \mathrm{T}_h J_h \mathrm{A} \Lambda(0)^{-1} \mathrm{A} y^0 .$$

and in the fully discrete case in (2.29)

$$y_h^0 = \mathrm{T}_h \Lambda_h^1 \mathrm{T}_h J_h \widehat{y}^0, \qquad \text{with} \qquad \widehat{y}^0 = \mathrm{A}(\Lambda^1)^{-1} \mathrm{A} y^0 . \tag{5.17}$$

The aim of this section is to prove the following bounds.

Lemma 5.5. *Under the assumptions of Theorem 2.8 it holds for* $\ell \in \{1,2,3\}$

$$\left\| \partial_t^\ell e_h(0) \right\|_X \leq C h^k .$$

Lemma 5.6. *Under the assumptions of Theorem 2.9 it holds for* $\ell \in \{1,2,3\}$

$$\left\| \partial_\tau^\ell e_h^\ell \right\|_X \leq C\big(\tau + h^k\big).$$

In order to conclude the desired assertion, we proceed in a series of lemmas and introduce the notation

$$\mathcal{R}^n = \big(\Lambda^n - \tau \mathrm{A}\big)^{-1}, \qquad \mathcal{R}_h^n = \big(\Lambda_h^n - \tau \mathrm{A}_h\big)^{-1}. \tag{5.18}$$

We provide a detailed proof of Lemma 5.6 first and explain afterwards how to conclude the assertion of Lemma 5.5. In order to keep the notation more simple, we assume without loss of generality in the following that the spatial order satisfies

$$k^* \leq \ell_{\max} - 2 , \tag{5.19}$$

with $\ell_{\max}$ defined in Assumption 2.1. Note that this directly implies the condition $\ell_{\max} \geq 4$. Hence, by (2.5) and (5.17)

$$\widehat{y}^0 = \mathrm{A}(\Lambda^1)^{-1} \mathrm{A} y^0 \in \mathcal{D}(\mathrm{A}^{k^*+1}). \tag{5.20}$$

In a first step we find a suitable representation for the discrete and exact solutions.

Lemma 5.7. *The numerical solution can be expanded via*

$$\begin{aligned}
\partial_\tau y_h^n &= \mathcal{R}_h^n \mathrm{A}_h y_h^{n-1} + G_h^n, \qquad G_h^n = \mathcal{R}_h^n \Lambda_h^n F_h^n\,,\\
\partial_\tau^2 y_h^n &= \mathcal{R}_h^n \mathrm{A}_h \mathcal{R}_h^{n-1} \mathrm{A}_h y_h^{n-2} + \big(\partial_\tau \mathcal{R}_h^n\big) \mathrm{A}_h y_h^{n-2} + \mathcal{R}_h^n \mathrm{A}_h G_h^{n-1} + \partial_\tau G_h^n\,,\\
\partial_\tau^3 y_h^n &= \mathcal{R}_h^n \mathrm{A}_h \mathcal{R}_h^{n-1} \mathrm{A}_h \mathcal{R}_h^{n-2} \mathrm{A}_h y_h^{n-3} + 2\partial_\tau \mathcal{R}_h^n \mathrm{A}_h \mathcal{R}_h^{n-2} \mathrm{A}_h y_h^{n-3}\\
&\quad + \mathcal{R}_h^n \mathrm{A}_h \partial_\tau \mathcal{R}_h^{n-1} \mathrm{A}_h y_h^{n-3} + \partial_\tau^2 \mathcal{R}_h^n \mathrm{A}_h y_h^{n-3}\\
&\quad + \mathcal{R}_h^n \mathrm{A}_h \mathcal{R}_h^{n-1} \mathrm{A}_h G_h^{n-2} + 2\partial_\tau \mathcal{R}_h^n \mathrm{A}_h G_h^{n-2} + \mathcal{R}_h^n \mathrm{A}_h \partial_\tau G_h^{n-1} + \partial_\tau^2 G_h^n\,.
\end{aligned}$$

The same holds for $\partial_\tau^\ell y(t^n)$, $\ell = 1,2,3$, with h formally set to zero and

$$G_h^n \to G^n = \mathcal{R}^n \Lambda^n \big(F^n + \delta^n\big), \qquad \delta^n = \partial_\tau y(t^n) - \partial_t y(t^n).$$

Proof. Starting from (2.27), we multiply by Λ_h^n and reorder the terms to obtain

$$\big(\Lambda_h^n - \tau \mathrm{A}_h\big) y_h^n = \Lambda_h^n y_h^{n-1} + \tau \Lambda_h^n F_h^n.$$

Using the resolvent, this further gives

$$y_h^n = y_h^{n-1} + \tau \mathcal{R}_h^n \mathrm{A}_h y_h^{n-1} + \tau \mathcal{R}_h^n \Lambda_h^n F_h^n,$$

which implies the representation for $\partial_\tau y_h^n$. The remaining identities are deduced from the product rule in Lemma 4.1. The results for the exact solution can be derived from the representation

$$\Lambda^n \partial_\tau y(t^n) = \mathrm{A} y(t^n) + \Lambda^n F^n + \Lambda^n \delta^n$$

and the same computations as above. □

In order to bound the expressions in Lemma 5.6, we subtract the representations for the exact and the numerical solution. Let us for example consider the first term in $\partial_\tau^3 e_h^3 = J_h \partial_\tau^3 y(t^3) - \partial_\tau^3 y_h^3$ given by

$$\begin{aligned}
\big(\partial_\tau^3 e_h^3\big)_1 &= J_h \mathcal{R}^3 \mathrm{A} \mathcal{R}^2 \mathrm{A} \mathcal{R}^1 \mathrm{A} y^0 - \mathcal{R}_h^3 \mathrm{A}_h \mathcal{R}_h^2 \mathrm{A}_h \mathcal{R}_h^1 \mathrm{A}_h y_h^0\\
&= \Big(J_h \mathcal{R}^3 \mathrm{A} \mathcal{R}^2 \mathrm{A} \mathcal{R}^1 \Lambda^1 \mathrm{A}^{-1} - \mathcal{R}_h^3 \mathrm{A}_h \mathcal{R}_h^2 \mathrm{A}_h \mathcal{R}_h^1 \Lambda_h^1 \mathrm{A}_h^{-1} J_h \Big) \widehat{y}^0,
\end{aligned}$$

where we used the initial value (2.29) with $\widehat{y}^0 = \mathrm{A}(\Lambda^1)^{-1} \mathrm{A} y^0$. In order to bound the difference, we proceed in two steps. First, we move the operators A and A_h to the right. Therefore, we employ the identities

$$\mathrm{A} \mathcal{R}^n \Lambda^n = \Lambda^n \mathcal{R}^n \mathrm{A}, \qquad \mathrm{A} \Lambda^n = \widetilde{\Lambda^n} \mathrm{A}, \tag{5.21}$$

with $\widetilde{\Lambda}$ defined in (5.11). The corresponding equalities are also valid for the discrete objects and the inverse $\mathrm{A}(\Lambda^n)^{-1} = (\widetilde{\Lambda^n})^{-1} \mathrm{A}$. Hence, we can write

$$J_h \mathcal{R}^3 \mathrm{A} \mathcal{R}^2 \mathrm{A} \mathcal{R}^1 \Lambda^1 \mathrm{A}^{-1} \widehat{y}^0 = J_h \mathcal{R}^3 \Lambda^2 \mathcal{R}^2 (\widetilde{\Lambda^2})^{-1} \widetilde{\Lambda^1} \Lambda^1 \mathcal{R}^1 (\widetilde{\Lambda^1})^{-1} \mathrm{A} \widehat{y}^0$$

and similarly for the discrete counterpart. A reformulation of Lemma 5.7 according to the above strategy is given in Lemma A.1. The differences of A and A_h as well as F and F_h are bounded by (5.15) and (5.16), respectively. The remaining differences in front of them are treated by the following abstract estimate. As a shorthand notation, we set

$$\prod_{j=1}^m B^j := B^m \dots B^1,\ m \geq 1, \qquad \text{and} \qquad \prod_{j=1}^m B^j := \mathrm{Id},\ m < 1.$$

Lemma 5.8. *Let $\mathcal{Y} \subseteq X$ be a Hilbert space, $m \in \mathbb{N}$, and consider operators $B_h^j \in \mathcal{L}(X_h)$ and $B^j \in \mathcal{L}(X)$, $j = 1, \dots, m$, with the following properties:*

$$\begin{aligned} \left\| \left(J_h B^j - B_h^j J_h\right) x \right\|_{X_h} &\lesssim \left(\tau + h^k\right) \|x\|_{\mathcal{Y}}, \\ \left\| B^j x \right\|_{\mathcal{Y}} &\lesssim \|x\|_{\mathcal{Y}}. \end{aligned} \tag{5.22}$$

Then, the product is bounded by

$$\left\| \left(J_h \left(\prod_{j=1}^m B^j \right) - \left(\prod_{j=1}^m B_h^j \right) J_h \right) x \right\|_{X_h} \lesssim \left(\tau + h^k\right) \|x\|_{\mathcal{Y}}.$$

Proof. Using the telescopic sum

$$\left(J_h \left(\prod_{j=1}^m B^j \right) - \left(\prod_{j=1}^m B_h^j \right) J_h \right) x = \sum_{\ell=1}^m \left(\prod_{j=\ell+1}^m B_h^j \right) \left(J_h B^\ell - B_h^\ell J_h\right) \left(\prod_{j=1}^{\ell-1} B^j \right) x,$$

we immediately conclude the assertion. □

With these preparations, we finally conclude the bounds for the initial values.

Proof of Lemma 5.6. We estimate the differences of the continuous and discretized operators in Lemma A.1. We split the proof in two parts. First, we compare the products of bounded operators involving

$$B^j \in \{\Lambda^n, \partial_\tau \Lambda^n, \partial_\tau^2 \Lambda^n, (\Lambda^n)^{-1}, \widetilde{\Lambda^n}, \partial_\tau \widetilde{\Lambda^n}, (\widetilde{\Lambda^n})^{-1}, \mathcal{R}^n, \mathrm{T}\} \tag{5.23}$$

and the discrete counterparts using Lemma 5.8 with $\mathcal{Y} = \mathcal{D}(A^{k^*})$. In the second step, we deal with the powers of A and A_h applied to the initial value, the right-hand side and the defects. Because of $\partial\Omega \in C^{k+1,1}$, we several times use the embedding $\mathcal{D}(\mathrm{A}^k) \hookrightarrow H^{k+1}(\Omega) \times H^k(\Omega)$, see, e.g., [17, Rem. 2.5.1.2].

(i) We first consider the operators involving Λ. Under Assumption 2.1 for $\kappa = k+1$ and with Lemmas 2.3 and 2.6 and the bound (5.12), the properties (5.22) are satisfied. For the resolvent $\mathcal{R}^n$, we directly employ Lemma A.2 to compute with (2.20) and (5.16)

$$\begin{aligned} \left\| \left(J_h \mathcal{R} - \mathcal{R}_h J_h\right) y \right\|_{X_h} &= \tau \left\| \mathcal{R}_h \left(J_h \mathrm{A} - \mathrm{A}_h J_h\right) \mathcal{R} y \right\|_{X_h} \\ &\leq C h^k \tau \left\| \mathrm{A}^{k^*+1} \mathcal{R} y \right\|_{X_h} \\ &\leq C h^k \left\| \mathrm{A}^{k^*} y \right\|_X. \end{aligned}$$

Finally using (5.13), T also satisfies (5.22) and Lemma 5.8 is applicable.

(ii) We employ Lemma A.1 and denote any product of operators from part (i) by Π and the discrete counterpart Π_h. Then, we have to compare expressions of the form

$$J_h \Pi (x + \delta) - \Pi_h x_h = (J_h \Pi - \Pi_h J_h) x + \Pi_h (J_h x - x_h) + J_h \Pi \delta$$

with

$$x \in \{\widehat{y}^0, \mathrm{A}\widehat{y}^0, \partial_\tau^\nu F^{\nu+1}, \mathrm{A}^\nu F^1\}, \quad \delta \in \{\partial_\tau^\nu \delta^{\nu+1}, \mathrm{A}^\nu \delta^1\} \tag{5.24}$$

and

$$x_h \in \{J_h \widehat{y}^0, \mathrm{A}_h J_h \widehat{y}^0, \partial_\tau^\nu F_h^{\nu+1}, \mathrm{A}_h^\nu F_h^1\}$$

where $\nu = 0, 1, 2$. Then, the first part is covered by part (i), provided that

$$\left\| \mathrm{A}^{k^*} x \right\|_X \leq C \tag{5.25}$$

holds for all x in (5.24). This follows from the assumptions of the theorem and (5.20). The second part is bounded due to (5.14), (5.15), and (5.16). Lastly, the estimate

$$\|J_h \Pi\delta\|_{X_h} \lesssim \left\|\mathrm{A}^2\delta\right\|_X$$

together with (5.9) yields

$$\|J_h \Pi\delta\|_{X_h} \lesssim \tau\big(\max_{\nu=0,1,2} \left\|\mathrm{A}^{\nu+2} y\right\|_{L^\infty(X)} + \max_{\nu=0,1,2} \left\|\mathrm{A}^2 \partial_\tau^\nu y\right\|_{L^\infty(X)}\big),$$

and hence the desired bound. □

Remark 5.9. We note that the bound in (5.25) is most restrictive for $x = \mathrm{A}\widehat{y}^0$, since it implies $\widehat{y}^0 \in \mathcal{D}(\mathrm{A}^{k^*+1})$. However, due to (2.5) one can at best achieve $\widehat{y}^0 \in \mathcal{D}(\mathrm{A}^{\ell_{\max}-1})$ which yields the restriction in (5.19).

We now prove Lemma 5.5 and observe that formally setting $\tau = 0$ in (2.27), we obtain the spatially discretized equation (2.23). This observation drives the following proof.

Proof of Lemma 5.5. Several steps in the proof simplify, and we mainly explain why we can weaken the assumption on $\ell_{\max}$. First note, that compared to the proof above we replace due to $\tau = 0$

$$\mathcal{R}^n \to (\Lambda^0)^{-1} \qquad\qquad \mathcal{R}_h^n \to (\Lambda_h^0)^{-1},$$

and the most delicate term is given by

$$\begin{aligned}\left(\partial_t^3 e_h(0)\right)_1 &= J_h(\Lambda^0)^{-1}\mathrm{A}(\Lambda^0)^{-1}\mathrm{A}(\Lambda^0)^{-1}\mathrm{A}y^0 - (\Lambda_h^0)^{-1}\mathrm{A}_h(\Lambda_h^0)^{-1}\mathrm{A}_h(\Lambda_h^0)^{-1}\mathrm{A}_h y_h^0 \\ &= \Big(J_h(\Lambda^0)^{-1}(\widetilde{\Lambda^0})^{-1}\mathrm{A} - (\Lambda_h^0)^{-1}(\widetilde{\Lambda_h^0})^{-1}\mathrm{A}_h J_h\Big)\mathrm{A}(\Lambda^0)^{-1}\mathrm{A}y^0,\end{aligned}$$

which yields the restriction $\ell_{\max} \geq 2$. Further, the terms in (5.23) reduce to

$$B^j \in \{\Lambda^0, \partial_t\Lambda^0, \partial_t^2\Lambda^0, (\Lambda^0)^{-1}, \widetilde{\Lambda^0}, \partial_t\widetilde{\Lambda^0}, (\widetilde{\Lambda^0})^{-1}, \mathrm{T}\}$$

and we employ Lemma 5.8 with $\mathcal{Y} = H^{k^*+1}(\Omega) \times H^{k^*}(\Omega)$. Hence, we eliminated all powers of A falling on $\widetilde{\Lambda}$ and no further restrictions on $\ell_{\max}$ enter and we proceed along the lines of Lemma 5.6. □

5.3. **Bounds on the first approximations.** This section is devoted to the proof of Theorem 2.12, i.e., we provide bounds for the first three approximations of the fully discrete scheme (2.27), which are not covered by Theorem 2.9. For the analysis, we rely on the following stability estimate for the resolvent as a self-mapping operator on $L^\infty(\Omega_h) \times L^\infty(\Omega_h)$.

Lemma 5.10. *Let Assumption 2.10 hold. Then there is a constant $C > 0$, which is independent of h, τ and n such that for all $\xi_h \in X_h$ the discrete resolvent $\mathcal{R}_h^n$ defined in (5.18) satisfies*

$$\|\mathcal{R}_h^n \xi_h\|_{L^\infty \times L^\infty} \leq C\frac{|\log(h)|}{\tau}\|\xi_h\|_{L^\infty \times L^\infty}.$$

Proof. Estimating the single components and employing Assumption 2.10 immediately yields the assertion. □

With this and the already derived bounds in the previous section, we finally present the proof of Theorem 2.12.

Proof of Theorem 2.12. As before, we decompose the error for $\ell \in \{0,1,2\}$ using (3.1) and estimate by (2.7), (2.9) and (2.12)

$$\begin{aligned}\big\|y(t^\ell) - \boldsymbol{\mathcal{L}_h} y_h^\ell\big\|_{L^\infty\times L^\infty} &\le \big\|\big(\mathrm{Id} - \boldsymbol{\mathcal{L}_h} J_h\big) y(t^\ell)\big\|_{L^\infty\times L^\infty} + \big\|e_h^\ell\big\|_{L^\infty\times L^\infty} \\ &\le Ch^k + \big\|e_h^\ell\big\|_{L^\infty\times L^\infty} \,.\end{aligned}$$

Hence, we only have to prove the bound

$$\big\|e_h^\ell\big\|_{L^\infty\times L^\infty} \le C\big(\tau + h^k\big)|\log(h)|^\ell \,,$$

which is done separately for $\ell = 0,1,2$.

(i) For $\ell = 0$, we insert the initial value (2.24) of the semi discretization

$$\big\|e_h^0\big\|_{L^\infty\times L^\infty} \le \big\|J_h y(t^0) - y_h(t^0)\big\|_{L^\infty\times L^\infty} + \big\|y_h(t^0) - y_h^0\big\|_{L^\infty\times L^\infty}$$

and estimate both parts. The first part satisfies due to Theorem 2.8

$$\big\|J_h y(t^0) - y_h(t^0)\big\|_{L^\infty\times L^\infty} \le Ch^k.$$

For the second part, we use the definitions of the initial values in (2.24) and (2.29) together with (5.21) to obtain

$$y_h(t^0) - y_h^0 = \mathrm{T}_h\Big(\Lambda_h^0 \mathrm{T}_h J_h (\widetilde{\Lambda^0})^{-1} - \Lambda_h^1 \mathrm{T}_h J_h (\widetilde{\Lambda^1})^{-1}\Big)\mathrm{A}^2 y^0.$$

Due to the stability of T_h, Λ_h, J_h, and Λ in $W^{1,\infty}\times W^{1,\infty}$, cf. Lemmas 2.3, 2.6 and 2.7 as well as (2.16), and with the Lipschitz continuity of Λ and Λ_h provided in Assumption 2.1, we infer

$$\big\|y_h(t^0) - y_h^0\big\|_{L^\infty\times L^\infty} \le C\tau\,\big\|\mathrm{A}^2 y^0\big\|_{W^{1,\infty},W^{1,\infty}} \,.$$

(ii) For $\ell = 1$, we expand

$$e_h^1 = e_h^0 + \tau\partial_\tau e_h^1 = e_h^0 + \tau J_h \partial_\tau y(t^1) - \tau\partial_\tau y_h^1 \,.$$

Since the first term is bounded by part (i) and the estimate for the second term follows from (2.16) and the regularity of $y \in C^1\big([0,T], W^{1,\infty}(\Omega)\times W^{1,\infty}(\Omega)\big)$, it is sufficient to provide a bound on $\partial_\tau y_h^1$. To this end, Lemma 5.7, the definition (5.17) of the initial value y_h^0, and the representation (2.25) of F_h yield

$$\partial_\tau y_h^1 = \mathcal{R}_h^1 \Lambda_h^1 \mathrm{T}_h J_h\big(\widehat{y}^0 + \mathrm{A}F^1\big) \,.$$

Thus, the identity

$$\mathcal{R}_h^1 \Lambda_h^1 \mathrm{T}_h = \mathrm{T}_h + \tau\mathcal{R}_h^1 \tag{5.26}$$

together with Lemmas 2.7 and 5.10 yields

$$\big\|\partial_\tau y_h^1\big\|_{L^\infty\times L^\infty} \le C|\log(h)|\big(\big\|\mathrm{A}^2 y^0\big\|_{W^{1,\infty}\times W^{1,\infty}} + \|f\|_{W^{1,\infty}}\big) \,.$$

(iii) For $\ell = 2$, we similarly write

$$e_h^2 = 2e_h^1 - e_h^0 + \tau^2\partial_\tau^2 e_h^2 = 2e_h^1 - e_h^0 + \tau^2 J_h \partial_\tau^2 y(t^2) - \tau\Big(\tau\partial_\tau^2 y_h^2\Big)$$

and only provide the missing bound on $\tau\partial_\tau^2 y_h^2$. Using the identity (5.21) we obtain from Lemma 5.7 together with (5.17) and (2.25)

$$\begin{aligned}\partial_\tau^2 y_h^2 &= \mathcal{R}_h^2 \Lambda_h^1 \mathcal{R}_h^1 J_h\big(\widehat{y}^0 + \mathrm{A}F^1\big) + \partial_\tau \mathcal{R}_h^2 \Lambda_h^1 \mathrm{T}_h J_h\big(\widehat{y}^0 + \mathrm{A}F^1\big) \\ &\quad + \mathcal{R}_h^2 \partial_\tau \Lambda_h^2 \mathrm{T}_h J_h \mathrm{A}F^1 + \mathcal{R}_h^2 \Lambda_h^2 \mathrm{T}_h J_h \mathrm{A}\partial_\tau F^2 \,.\end{aligned} \tag{5.27}$$

Employing the property (5.26) several times, we obtain

$$\left\|\tau\partial_\tau^2 y_h^2\right\|_{L^\infty\times L^\infty} \le C|\log(h)|^2\big(\|y\|_{H^6\times W^{3,\infty}} + \|f\|_{H^4}\big),$$

see Lemma A.3 for details. □

5.4. **Adaptions for Corollary 2.13.** Let us briefly discuss how the above results can be improved if it is sufficient to have maximum norm convergence only for u. The proofs of Sections 3 and 4 can be repeated simply using one derivative less. Also the defects require less regularity in time, but can be bounded similar to Lemmas 5.2 and 5.4.

However, the theory changes slightly when it comes to the initial values. In the semi discrete case, we differentiate the error only twice, and hence need to bound similar to Lemma 5.5

$$\begin{aligned}\big(\partial_t^2 e_h(0)\big)_1 &= J_h(\Lambda^0)^{-1}\mathrm{A}(\Lambda^0)^{-1}\mathrm{A}y^0 - (\Lambda_h^0)^{-1}\mathrm{A}_h(\Lambda_h^0)^{-1}\mathrm{A}_h y_h^0\\ &= \Big(J_h(\Lambda^0)^{-1}(\widetilde{\Lambda^0})^{-1}\mathrm{A} - (\Lambda_h^0)^{-1}(\widetilde{\Lambda_h^0})^{-1}\mathrm{A}_h J_h\Big)\mathrm{A}y^0\,,\end{aligned}$$

which only yields the restriction $\ell_{\max} \ge 1$, and we proceed as in Lemma 5.5.

For the two discrete derivatives of the error, we have to prove Lemma 5.6 only for $\ell = 1, 2$. The most critical term is then by (2.31)

$$\begin{aligned}\big(\partial_\tau^2 e_h^2\big)_1 &= J_h\mathcal{R}^2\mathrm{A}\mathcal{R}^1\mathrm{A}y^0 - \mathcal{R}_h^2\mathrm{A}_h\mathcal{R}_h^1\mathrm{A}_h y_h^0\\ &= \Big(J_h\mathcal{R}^2\mathrm{A}\mathcal{R}^1 - \mathcal{R}_h^2\mathrm{A}_h\mathcal{R}_h^1 J_h\Big)\mathrm{A}y^0\\ &= \Big(J_h\mathcal{R}^2\Lambda^1\mathcal{R}^1(\widetilde{\Lambda^1})^{-1}\mathrm{A} - \mathcal{R}_h^2\Lambda_h^1\mathcal{R}_h^1(\widetilde{\Lambda_h^1})^{-1}\mathrm{A}_h J_h\Big)\mathrm{A}y^0.\end{aligned}$$

We note that in fact Lemma 5.8 is also applicable under the condition $k^* \le \ell_{\max} - 1$. However, this is sufficient as there does not enter another restriction on k^* from the choice of the initial value y_h^0, cf. Remark 5.9. The rest of the proof then is along the lines of Lemma 5.6.

Acknowledgments

The authors thank Roland Schnaubelt and Konstantin Zerulla for helpful discussions on elliptic regularity and in particular Konstantin Zerulla and Rebekka Zimmermann for their careful reading of this manuscript.

References

[1] P. F. Antonietti, I. Mazzieri, M. Muhr, V. Nikolić, and B. Wohlmuth, *A high-order discontinuous Galerkin method for nonlinear sound waves*, J. Comput. Phys. **415** (2020), 109484, 27. MR4093710

[2] W. Arendt, C. J. K. Batty, M. Hieber, and F. Neubrander, *Vector-valued Laplace transforms and Cauchy problems*, Second, Monographs in Mathematics, vol. 96, Birkhäuser/Springer Basel AG, Basel, 2011. MR2798103

[3] N. Yu. Bakaev, *Maximum norm resolvent estimates for elliptic finite element operators*, BIT **41** (2001), no. 2, 215–239. MR1837395

[4] N. Yu. Bakaev, V. Thomée, and L. B. Wahlbin, *Maximum-norm estimates for resolvents of elliptic finite element operators*, Math. Comp. **72** (2003), no. 244, 1597–1610. MR1986795

[5] G. A. Baker, H. H. Bramble, and V. Thomée, *Single step Galerkin approximations for parabolic problems*, Math. Comp. **31** (1977), no. 140, 818–847. MR448947

[6] G. A. Baker and V. A. Dougalis, *On the L^∞-convergence of Galerkin approximations for second-order hyperbolic equations*, Math. Comp. **34** (1980), no. 150, 401–424. MR559193

[7] G. A. Baker, V. A. Dougalis, and S. M. Serbin, *High order accurate two-step approximations for hyperbolic equations*, RAIRO Anal. Numér. **13** (1979), no. 3, 201–226. MR543933
[8] L. A. Bales and V. A. Dougalis, *Cosine methods for nonlinear second-order hyperbolic equations*, Math. Comp. **52** (1989), no. 186, 299–319, S15–S33. MR955747
[9] L. A. Bales, V. A. Dougalis, and S. M. Serbin, *Cosine methods for second-order hyperbolic equations with time-dependent coefficients*, Math. Comp. **45** (1985), no. 171, 65–89. MR790645
[10] J. H. Bramble, A. H. Schatz, V. Thomée, and L. B. Wahlbin, *Some convergence estimates for semidiscrete Galerkin type approximations for parabolic equations*, SIAM J. Numer. Anal. **14** (1977), no. 2, 218–241. MR448926
[11] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods*, Third, Texts in Applied Mathematics, vol. 15, Springer, New York, 2008. MR2373954 (2008m:65001)
[12] P. Chatzipantelidis, R. D. Lazarov, V. Thomée, and L. B. Wahlbin, *Parabolic finite element equations in nonconvex polygonal domains*, BIT **46** (2006), no. suppl., S113–S143. MR2283311
[13] B. Dörich, J. Leibold, and B. Maier, *Stability of the Ritz map in the maximum norm for non-conforming finite elements*, Karlsruhe Institute of Technology, 2021. https://doi.org/10.5445/IR/1000141502.
[14] C. M. Elliott and T. Ranner, *Finite element analysis for a coupled bulk-surface partial differential equation*, IMA J. Numer. Anal. **33** (2013), no. 2, 377–402. MR3047936
[15] ———, *A unified theory for continuous-in-time evolving finite element space approximations to partial differential equations in evolving domains*, IMA J. Numer. Anal. **41** (2021), no. 3, 1696–1845. MR4286249
[16] T. Geveci, *On the convergence of Galerkin approximation schemes for second-order hyperbolic equations in energy and negative norms*, Math. Comp. **42** (1984), no. 166, 393–415. MR736443
[17] P. Grisvard, *Elliptic problems in nonsmooth domains*, Monographs and Studies in Mathematics, vol. 24, Pitman (Advanced Publishing Program), Boston, MA, 1985. MR775683
[18] D. Hipp, M. Hochbruck, and C. Stohrer, *Unified error analysis for nonconforming space discretizations of wave-type equations*, IMA J. Numer. Anal. **39** (2019), no. 3, 1206–1245. MR3984056
[19] M. Hochbruck and B. Maier, *Error analysis for space discretizations of quasilinear wave-type equations*, IMA J. Numer. Anal. (2021). https://doi.org/10.1093/imanum/drab073.
[20] M. Hochbruck and T. Pažur, *Error analysis of implicit Euler methods for quasilinear hyperbolic evolution equations*, Numer. Math. **135** (2017), no. 2, 547–569. MR3599569
[21] T. Kashiwabara and T. Kemmochi, *Pointwise error estimates of linear finite element method for Neumann boundary value problems in a smooth domain*, Numer. Math. **144** (2020), no. 3, 553–584. MR4071825
[22] B. Kok and T. Geveci, *The convergence of Galerkin approximation schemes for second-order hyperbolic equations with dissipation*, Math. Comp. **44** (1985), no. 170, 379–390, S17–S25. MR777270
[23] B. Kovács and C. A. Power Guerra, *Maximum norm stability and error estimates for the evolving surface finite element method*, Numer. Methods Partial Differential Equations **34** (2018), no. 2, 518–554. MR3765711
[24] D. Leykekhman and B. Vexler, *Finite element pointwise results on convex polyhedral domains*, SIAM J. Numer. Anal. **54** (2016), no. 2, 561–587. MR3470741
[25] H. Liao and Z. Sun, *Maximum norm error estimates of efficient difference schemes for second-order wave equations*, J. Comput. Appl. Math. **235** (2011), no. 8, 2217–2233. MR2763137
[26] P. Linz, *Numerical methods for Volterra integral equations of the first kind*, Comput. J. **12** (1969), 393–397. MR253577
[27] W. Littman, *The wave operator and L_p norms*, J. Math. Mech. **12** (1963), 55–68. MR0146521
[28] A. Lunardi, *Interpolation theory*, Appunti. Scuola Normale Superiore di Pisa (Nuova Serie) [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)], vol. 16, Edizioni della Normale, Pisa, 2018. Third edition [of MR2523200]. MR3753604
[29] B. Maier, *Error analysis for space and time discretizations of quasilinear wave-type equations*, Ph.D. Thesis, 2020. https://doi.org/10.5445/IR/1000120935.

[30] ———, *Error analysis for full discretizations of quasilinear wave-type equations with two variants of the implicit midpoint rule*, Technical Report 2021/24, Karlsruhe Institute of Technology, 2021. https://doi.org/10.5445/IR/1000133244.
[31] C. G. Makridakis, *Finite element approximations of nonlinear elastic waves*, Math. Comp. **61** (1993), no. 204, 569–594. MR1195426
[32] F. Natterer, *Über die punktweise Konvergenz finiter Elemente*, Numer. Math. **25** (1975/76), no. 1, 67–77. MR474884
[33] J. A. Nitsche, *L_∞-convergence of finite element approximation*, Journées "Éléments Finis" (Rennes, 1975), 1975, pp. 18. MR568857
[34] ———, *L_∞-convergence of finite element approximations*, Mathematical aspects of finite element methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975), 1977, pp. 261–274. Lecture Notes in Math., Vol. 606. MR0488848
[35] C. Palencia, *Maximum norm analysis of completely discrete finite element methods for parabolic problems*, SIAM J. Numer. Anal. **33** (1996), no. 4, 1654–1668. MR1403564
[36] R. Rannacher, *Zur L^∞-Konvergenz linearer finiter Elemente beim Dirichlet-Problem*, Math. Z. **149** (1976), no. 1, 69–77. MR488859
[37] R. Rannacher and R. Scott, *Some optimal error estimates for piecewise linear finite element approximations*, Math. Comp. **38** (1982), no. 158, 437–445. MR645661
[38] A. H. Schatz, V. Thomée, and L. B. Wahlbin, *Stability, analyticity, and almost best approximation in maximum norm for parabolic finite element equations*, Comm. Pure Appl. Math. **51** (1998), no. 11-12, 1349–1385. MR1639143
[39] A. H. Schatz, V. C. Thomée, and L. B. Wahlbin, *Maximum norm stability and error estimates in parabolic finite element equations*, Comm. Pure Appl. Math. **33** (1980), no. 3, 265–304. MR562737
[40] A. H. Schatz and L. B. Wahlbin, *Interior maximum norm estimates for finite element methods*, Math. Comp. **31** (1977), no. 138, 414–442. MR431753
[41] ———, *On the quasi-optimality in L_∞ of the $\dot{H}^1$-projection into finite element spaces*, Math. Comp. **38** (1982), no. 157, 1–22. MR637283
[42] ———, *Interior maximum-norm estimates for finite element methods. II*, Math. Comp. **64** (1995), no. 211, 907–928. MR1297478
[43] R. Scott, *Optimal L^∞ estimates for the finite element method on irregular meshes*, Math. Comp. **30** (1976), no. 136, 681–697. MR436617
[44] V. Thomée and L. B. Wahlbin, *Maximum-norm estimates for finite-element methods for a strongly damped wave equation*, BIT **44** (2004), no. 1, 165–179. MR2057368
[45] L. B. Wahlbin, *Maximum norm error estimates in the finite element method with isoparametric quadratic elements and numerical integration*, RAIRO Anal. Numér. **12** (1978), no. 2, 173–202, v. MR502070
[46] M. Zlámal, *On the finite element method*, Numer. Math. **12** (1968), 394–409. MR243753

Appendix A.

In this appendix we provide the proof of Lemma 2.7 and the postponed calculations from Section 5. The proof is adapted from the conforming case presented in [10, Lem. 4.1].

Proof of Lemma 2.7. Let $q \geq 1$ and $\xi_h = (\varphi_h, \psi_h)^T \in X_h$. Note that it is sufficient to prove

$$\|S_h\psi_h\|_{L^p} \leq \|\psi_h\|_{L^r}\,, \tag{A.1}$$

for $p \geq 2$ and $r \geq 1$ with $0 \leq \frac{1}{r} - \frac{1}{p} < \frac{1}{N}$.

For $p = 2$ and $r \geq 1$ with $\frac{1}{2} \leq \frac{1}{r} < \frac{1}{N} + \frac{1}{2}$, we have

$$\|S_h\psi_h\|_{V_h}^2 = (S_h\psi_h \mid S_h\psi_h)_{V_h} = (\psi_h \mid S_h\psi_h)_{H_h} \lesssim \|\psi_h\|_{L^r} \|S_h\psi_h\|_{L^{\frac{r}{r-1}}}\,.$$

Due to $\frac{r}{r-1} < \infty$ for $N = 2$ and $\frac{r}{r-1} < \frac{2N}{N-2}$ for $N = 3$, Sobolev's embedding, and Poincaré's inequality, imply the bound

$$\|S_h\psi_h\|_{L^2} \lesssim \|\psi_h\|_{L^r}\,.$$

For $p = \infty$, we define the modified solution operator

$$\widetilde{S}_h = \mathcal{L}_h^{V*} S \mathcal{L}_h,$$

which satisfies by Assumption 2.5, Theorem 2.4 for $r > N$, and (2.7)

$$\big\|\widetilde{S}_h\psi_h\big\|_{L^\infty(\Omega_h)} \lesssim \big\|S\mathcal{L}_h\psi_h\big\|_{W^{1,\infty}(\Omega)} \lesssim \big\|S\mathcal{L}_h\psi_h\big\|_{W^{2,r}(\Omega)} \lesssim \big\|\psi_h\big\|_{L^r(\Omega_h)}.$$

It remains to bound the difference by the inverse estimate (2.15)

$$\begin{aligned}
\big\|\widetilde{S}_h\psi_h - S_h\psi_h\big\|_{L^\infty(\Omega_h)} &\lesssim h^{-N/6}\big\|\widetilde{S}_h\psi_h - S_h\psi_h\big\|_{V_h} \\
&= h^{-N/6} \sup_{\|\varphi_h\|_{V_h}=1} \Big(\widetilde{S}_h\psi_h - S_h\psi_h \mid \varphi_h\Big)_{V_h} \\
&= h^{-N/6} \sup_{\|\varphi_h\|_{V_h}=1} \big((\psi_h \mid \varphi_h)_{H_h} - (\mathcal{L}_h\psi_h \mid \mathcal{L}_h\varphi_h)_H\big).
\end{aligned} \tag{A.2}$$

We use a variant of [15, Lem. 8.24] to obtain

$$\big|(\psi_h \mid \varphi_h)_{H_h} - (\mathcal{L}_h\psi_h \mid \mathcal{L}_h\varphi_h)_H\big| \lesssim h\,\|\mathcal{L}_h\psi_h\|_{L^r}\,\|\mathcal{L}_h\psi_h\|_{L^{r'}} \lesssim h\,\|\psi_h\|_{L^r}\,\|\psi_h\|_{V_h}\,,$$

and (A.1) follows for $p = \infty$. Thus, an interpolation argument yields (A.1) for all $2 \le p \le \infty$, see, e.g., [28, Thm. 2.6]. □

Next, we give an extension of Lemma 5.7. We do not provide a proof here, since the expressions are derived by an iterative application of the identities (5.21).

Lemma A.1. *Let $\widehat{y}^0$ be given by (5.20). Then, it holds*

$$\begin{aligned}
\partial_\tau y_h^1 &= \mathcal{R}_h^1\Lambda_h^1\big(\mathrm{T}_h J_h\widehat{y}^0 + F_h^1\big)\,,\\
\partial_\tau^2 y_h^2 &= \mathcal{R}_h^2\Lambda_h^1\mathcal{R}_h^1\big(J_h\widehat{y}^0 + \mathrm{A}_h F_h^1\big) - \mathcal{R}_h^2\partial_\tau\Lambda_h^2\mathcal{R}_h^1\Lambda_h^1\big(\mathrm{T}_h J_h\widehat{y}^0 + F_h^1\big)\\
&\quad + \mathcal{R}_h^2\Lambda_h^2\partial_\tau F_h^2 + \mathcal{R}_h^2\partial_\tau\Lambda_h^2 F_h^1\,,\\
\partial_\tau^3 y_h^3 &= \mathcal{R}_h^3\Lambda_h^2\mathcal{R}_h^2(\widetilde{\Lambda_h^2})^{-1}\widetilde{\Lambda_h^1}\Lambda_h^1\mathcal{R}_h^1(\widetilde{\Lambda_h^1})^{-1}\mathrm{A}_h\big(J_h\widehat{y}^0 + \mathrm{A}_h F_h^1\big)\\
&\quad - 2\mathcal{R}_h^3\partial_\tau\Lambda_h^3\mathcal{R}_h^2\Lambda_h^1\mathcal{R}_h^1\big(J_h\widehat{y}^0 + \mathrm{A}_h F_h^1\big) - \mathcal{R}_h^3\Lambda_h^2\mathcal{R}_h^2(\widetilde{\Lambda_h^2})^{-1}\partial_\tau\widetilde{\Lambda_h^2}\Lambda_h^1\mathcal{R}_h^1 J_h\widehat{y}^0\\
&\quad + \big(2\mathcal{R}_h^3\partial_\tau\Lambda_h^3\mathcal{R}_h^2\partial_\tau\Lambda_h^2\mathcal{R}_h^1 - \mathcal{R}_h^3\partial_\tau^2\Lambda_h^3\mathcal{R}_h^1\big)\Lambda_h^1\mathrm{T}_h J_h\widehat{y}^0\\
&\quad + \mathcal{R}_h^3\Lambda_h^2\mathcal{R}_h^2\mathrm{A}_h\partial_\tau F_h^2 + \mathcal{R}_h^3\Lambda_h^2\mathcal{R}_h^2(\widetilde{\Lambda_h^2})^{-1}\partial_\tau\widetilde{\Lambda_h^2}\mathrm{A}_h F_h^1\\
&\quad - \mathcal{R}_h^3\Lambda_h^2\mathcal{R}_h^2(\widetilde{\Lambda_h^2})^{-1}\partial_\tau\widetilde{\Lambda_h^2}\Lambda_h^1\mathcal{R}_h^1\mathrm{A}_h F_h^1\\
&\quad + \mathcal{R}_h^3\Lambda_h^3\partial_\tau^2 F_h^3 + \mathcal{R}_h^3\partial_\tau\Lambda_h^3\partial_\tau F_h^2 - \mathcal{R}_h^3\partial_\tau\Lambda_h^3\mathcal{R}_h^2\Lambda_h^2\partial_\tau F_h^2\\
&\quad + \mathcal{R}_h^3\partial_\tau\Lambda_h^3\partial_\tau F_h^2 + \mathcal{R}_h^3\partial_\tau^2\Lambda_h^3 F_h^1 - \mathcal{R}_h^3\partial_\tau\Lambda_h^3\mathcal{R}_h^2\partial_\tau\Lambda_h^2 F_h^1\\
&\quad - \mathcal{R}_h^3\partial_\tau\Lambda_h^3\mathcal{R}_h^2\Lambda_h^2\partial_\tau F_h^2 - \mathcal{R}_h^3\partial_\tau\Lambda_h^3\mathcal{R}_h^2\partial_\tau\Lambda_h^2 F_h^1 + \mathcal{R}_h^3\partial_\tau\Lambda_h^3\mathcal{R}_h^2\partial_\tau\Lambda_h^2\mathcal{R}_h^1\Lambda_h^1 F_h^1\\
&\qquad - \mathcal{R}_h^3\partial_\tau^2\Lambda_h^3\mathcal{R}_h^1\Lambda_h^1 F_h^1 + \mathcal{R}_h^3\partial_\tau\Lambda_h^3\mathcal{R}_h^2\partial_\tau\Lambda_h^2\mathcal{R}_h^1\Lambda_h^1 F_h^1\,.
\end{aligned}$$

The same expansion holds for $\partial_\tau^\ell y(t^\ell)$, $\ell = 1, 2, 3$, with h formally set to zero and

$$F_h^n \to F^n + \delta^n, \qquad \delta^n = \partial_\tau y(t^n) - \partial_t y(t^n)\,.$$

Since, in the case $\Lambda \neq \mathrm{Id}$, the operator A does in general not commute with the resolvent, we provide the bounds which are still available given Assumption 2.1.

Lemma A.2. *Let $\mathcal{R}$ be the resolvent defined in (5.18) and ℓ_{max} given in Assumption 2.1. Then, for $0 \leq \ell \leq \ell_{max}$ there are constants C_ℓ such that*

$$\left\| \mathrm{A}^\ell \mathcal{R} y \right\|_X \leq C_\ell \left\| \mathrm{A}^\ell y \right\|_X$$
$$\tau \left\| \mathrm{A}^{\ell+1} \mathcal{R} y \right\|_X \leq C_\ell \left\| \mathrm{A}^\ell y \right\|_X .$$

for all $y \in \mathcal{D}(\mathrm{A}^{\ell_{max}})$.

Proof. By the skew-adjointness of A and (2.4), we derive

$$\|\mathcal{R} y\|_X \leq C_\lambda \|y\|_X ,$$

which is the first equation for $\ell = 0$. For $0 \leq \ell \leq \ell_{\max} - 1$, we compute

$$\begin{aligned} \left\| \mathrm{A}^{\ell+1} \mathcal{R} y \right\|_X &= \left\| \mathrm{A}^\ell \Lambda \mathcal{R} (\widetilde{\Lambda})^{-1} \mathrm{A} y \right\|_X \\ &\leq \|\Lambda\|_{\mathcal{D}(\mathrm{A}^\ell) \leftarrow \mathcal{D}(\mathrm{A}^\ell)} \|\mathcal{R}\|_{\mathcal{D}(\mathrm{A}^\ell) \leftarrow \mathcal{D}(\mathrm{A}^\ell)} \left\| (\widetilde{\Lambda})^{-1} \right\|_{\mathcal{D}(\mathrm{A}^\ell) \leftarrow \mathcal{D}(\mathrm{A}^\ell)} \left\| \mathrm{A}^{\ell+1} y \right\|_X \end{aligned}$$

and the claim follows by induction. The second estimate is due to the resolvent identity

$$\tau \mathrm{A}^{\ell+1} \mathcal{R} y = \mathrm{A}^\ell \Lambda \mathcal{R} y - \mathrm{A}^\ell y$$

and an application of the above bound. □

We finally prove the bound missing for the convergence of e_h^2.

Lemma A.3. *The discrete derivative in (5.27) satisfies the bound*

$$\left\| \tau \partial_\tau^2 y_h^2 \right\|_{L^\infty \times L^\infty} \leq C |\log(h)|^2 \big(\|y\|_{H^6 \times W^{3,\infty}} + \|f\|_{H^4} \big)$$

with a constant C independent of τ and h.

Proof. We decompose into four parts by

$$\begin{aligned} \partial_\tau^2 y_h^2 &= \mathcal{R}_h^2 \Lambda_h^1 \mathcal{R}_h^1 J_h \big(\widehat{y}^0 + \mathrm{A} F^1 \big) + \partial_\tau \mathcal{R}_h^2 \Lambda_h^1 \mathrm{T}_h J_h \big(\widehat{y}^0 + \mathrm{A} F^1 \big) \\ &\quad + \mathcal{R}_h^2 \partial_\tau \Lambda_h^2 \mathrm{T}_h J_h \mathrm{A} F^1 + \mathcal{R}_h^2 \Lambda_h^2 \mathrm{T}_h J_h \mathrm{A} \partial_\tau F^2 \\ &=: \sum_{i=1}^4 \big(\partial_\tau^2 y_h^2 \big)_i \end{aligned}$$

For the first term, we employ (5.26)

$$\begin{aligned} \tau \big(\partial_\tau^2 y_h^2 \big)_1 &= \tau \mathcal{R}_h^2 \Lambda_h^1 \mathcal{R}_h^1 J_h \big(\widehat{y}^0 + \mathrm{A} F^1 \big) \\ &= \tau \mathcal{R}_h^2 \mathrm{A}_h \mathcal{R}_h^1 \Lambda_h^1 \mathrm{T}_h J_h \big(\widehat{y}^0 + \mathrm{A} F^1 \big) \\ &= \tau \mathcal{R}_h^2 \big(\mathrm{Id} + \tau \Lambda_h^1 \mathcal{R}_h^1 \mathrm{A}_h (\Lambda_h^1)^{-1} \big) J_h \big(\widehat{y}^0 + \mathrm{A} F^1 \big) \\ &= \big(\tau \mathcal{R}_h^2 \big) J_h \big(\widehat{y}^0 + \mathrm{A} F^1 \big) + \big(\tau \mathcal{R}_h^2 \big) \Lambda_h^1 \big(\tau \mathcal{R}_h^1 \big) (\widetilde{\Lambda_h^1})^{-1} \mathrm{A}_h J_h \big(\widehat{y}^0 + \mathrm{A} F^1 \big) \\ &= \big(\tau \mathcal{R}_h^2 \big) J_h \big(\widehat{y}^0 + \mathrm{A} F^1 \big) + \big(\tau \mathcal{R}_h^2 \big) \Lambda_h^1 \big(\tau \mathcal{R}_h^1 \big) (\widetilde{\Lambda_h^1})^{-1} J_h \mathrm{A} \big(\widehat{y}^0 + \mathrm{A} F^1 \big) \\ &\quad + \big(\tau \mathcal{R}_h^2 \big) \Lambda_h^1 \big(\tau \mathcal{R}_h^1 \big) (\widetilde{\Lambda_h^1})^{-1} \big(\mathrm{A}_h J_h - J_h \mathrm{A} \big) \big(\widehat{y}^0 + \mathrm{A} F^1 \big), \end{aligned}$$

and use Lemmas 2.6 and 5.10 and the bounds (5.3) and (5.4). We further have by (5.26) that

$$\begin{aligned}\tau\big(\partial_\tau^2 y_h^2\big)_2 &= \tau\partial_\tau \mathcal{R}_h^2 \Lambda_h^1 \mathrm{T}_h J_h\big(\widehat{y}^0 + \mathrm{A}F^1\big)\\ &= -\big(\tau\mathcal{R}_h^2\big)\big(\partial_\tau\Lambda_h^2\big)\mathcal{R}_h^1\Lambda_h^1\mathrm{T}_h J_h\big(\widehat{y}^0 + \mathrm{A}F^1\big)\\ &= -\big(\tau\mathcal{R}_h^2\big)\big(\partial_\tau\Lambda_h^2\big)\big(\mathrm{T}_h + \tau\mathcal{R}_h^1\big)J_h\big(\widehat{y}^0 + \mathrm{A}F^1\big),\end{aligned}$$

and the bounds follow together with Lemma 2.7 as above. For the last two terms, we observe

$$\begin{aligned}\tau\big(\partial_\tau^2 y_h^2\big)_3 &= \big(\tau\mathcal{R}_h^2\big)\partial_\tau\Lambda_h^2\mathrm{T}_h J_h \mathrm{A}F^1\\ \tau\big(\partial_\tau^2 y_h^2\big)_4 &= \big(\tau\mathcal{R}_h^2\big)\Lambda_h^2\mathrm{T}_h J_h\mathrm{A}\partial_\tau F^2\end{aligned}$$

such that Lemmas 2.6 and 5.10 yield the assertion. □

Institute for Applied and Numerical Mathematics, Karlsruhe Institute of Technology, 76149 Karlsruhe, Germany

Email address: benjamin.doerich@kit.edu, jan.leibold@kit.edu