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

Benjamin Dörich, Jan Leibold, Bernhard Maier

CRC Preprint 2021/47 (revised), September 2023
Maximum norm error bounds for the full discretization of non-autonomous wave equations

BENJAMIN DÖRICH†, JAN LIEBOLD AND BERNHARD MAIER
Institute for Applied and Numerical Mathematics, Karlsruhe Institute of Technology, Englerstr. 2, 76131 Karlsruhe, Germany

[Received on 6 November 2023]

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 optimal error bounds in the $W_{1,\infty} \times L_\infty$-norm 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 (discrete) time derivatives on the error in the $H_1 \times L_2$-norm which are reduced to estimates of the differentiated initial errors. To confirm our theoretical findings, we also present numerical experiments.

Keywords: error analysis, full discretization, wave equation, maximum norm error bounds, nonconforming space discretization, isoparametric finite elements, a-priori error bounds

1. Introduction

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

\[
\partial_{tt} u(t, x) = \lambda(t, x) - \frac{1}{\Delta} u(t, x) + f(t, x), \quad t \in [0, T], \quad x \in \Omega,
\]

on the domain $\Omega \subseteq \mathbb{R}^N$, $N = 2, 3$. We assume it to be bounded and convex with a sufficiently regular boundary, and impose homogeneous Dirichlet boundary conditions and appropriate initial conditions. We discretize (1.1) with isoparametric finite elements in space and the implicit Euler scheme in time, and derive $W_{1,\infty} \times L_\infty$-norm error bounds both 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 $\Omega$. Compared to the classical estimates in $L_2$, see, e.g., Bales et al. (1985); Bales & Dougalis (1989), which are implied (with non-optimal order) by our maximum norm error estimates, and in the energy space $H_1$, see, e.g., Maier (2022); Hochbruck & Maier (2021), they provide an additional insight in the approximation quality. For example, they become particularly interesting if one wants to approximate the quasilinear wave equation

\[
\partial_{tt} u(t, x) = \lambda(u(t, x)) - \frac{1}{\Delta} u(t, x) + f(t, x, u(t, x)),
\]

where for example in nonlinear acoustics $\lambda(u) = 1 - u^m$, $m = 2, 3$. The reason is, that this equation is only well-posed as long as $\lambda(u)$ satisfies a pointwise lower bound away from zero. When discretizing (1.2) in space, it has to be ensured that the spatial discretization inherits this property. Since this requires

---

† Corresponding author. Email: benjamin.doerich@kit.edu

© The author 2023. Published by Oxford University Press on behalf of the Institute of Mathematics and its Applications. All rights reserved.
Maximum norm error bounds for the full discretization of non-autonomous wave equations

Benjamin Dörich†, Jan Leibold and Bernhard Maier

Institute for Applied and Numerical Mathematics, Karlsruhe Institute of Technology, Englerstr. 2, 76131 Karlsruhe, Germany

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 optimal error bounds in the $W^{1,\infty} \times L^\infty$-norm 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 (discrete) time derivatives on the error in the $H^1 \times L^2$-norm which are reduced to estimates of the differentiated initial errors. To confirm our theoretical findings, we also present numerical experiments.

Keywords: error analysis, full discretization, wave equation, maximum norm error bounds, nonconforming space discretization, isoparametric finite elements, a-priori error bounds

1. Introduction

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

$$\partial_t u(t, \mathbf{x}) = \lambda(t, \mathbf{x})^{-1} \Delta u(t, \mathbf{x}) + f(t, \mathbf{x}), \quad t \in [0, T], \mathbf{x} \in \Omega, \quad (1.1)$$

on the domain $\Omega \subseteq \mathbb{R}^N$, $N = 2, 3$. We assume it to be bounded and convex with a sufficiently regular boundary, and impose homogeneous Dirichlet boundary conditions and appropriate initial conditions. We discretize (1.1) with isoparametric finite elements in space and the implicit Euler scheme in time, and derive $W^{1,\infty} \times L^\infty$-norm error bounds both 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 $\Omega$. Compared to the classical estimates in $L^2$, see, e.g., Bales et al. (1985); Bales & Dougalis (1989), which are implied (with non-optimal order) by our maximum norm error estimates, and in the energy space $H^1$, see, e.g., Maier (2022); Hochbruck & Maier (2021), they provide an additional insight in the approximation quality. For example, they become particularly interesting if one wants to approximate the quasilinear wave equation

$$\partial_t u(t, \mathbf{x}) = \lambda(u(t, \mathbf{x}))^{-1} \Delta u(t, \mathbf{x}) + f(t, \mathbf{x}, u(t, \mathbf{x}), \partial_t u(t, \mathbf{x})), \quad (1.2)$$

where for example in nonlinear acoustics $\lambda(u) = 1 - u^m$, $m = 2, 3$. The reason is, that this equation is only well-posed as long as $\lambda(u)$ satisfies a pointwise lower bound away from zero. When discretizing (1.2) in space, it has to be ensured that the spatial discretization inherits this property. Since this requires

†Corresponding author. Email: benjamin.doerich@kit.edu
a pointwise bound of the numerical approximation, maximum norm estimates, as they are provided in this paper, are sufficient - once they are transferred to the nonlinear case - to guarantee such constraints. Indeed, for some spatial discretization $u_h$ the triangle inequality

$$||u_h||_{L^\infty} \leq ||u||_{L^\infty} + ||u - u_h||_{L^\infty}$$

allows to keep the maximum norm of the numerical solution arbitrarily close to the one of the exact solution once convergence is established. So far, to show convergence 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., Maier (2022); Antonietti et al. (2020); Makridakis (1993). Alternatively, $H^2$-conforming finite elements, as suggested in Zlámal (1968), 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) using Lagrangian finite element methods without CFL conditions and also generalizes to higher-order methods in time.

In the articles Baker et al. (1979); Baker & Dougalis (1980), 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 $W^{1,\infty} \times L^\infty$-norm, 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 $\Omega$. Unfortunately, this prevents us from using these results on a computational domain $\Omega_h$ with a piecewise polynomial boundary. Our research is mainly inspired by Baker et al. (1979); Baker & Dougalis (1980) and we are not aware of further maximum norm estimates for wave equations discretized by finite elements, besides the one-dimensional case considered in Trautmann et al. (2018). For finite differences on a square combined with a fourth-order in time scheme, an error bound under a CFL condition is established in Liao & Sun (2011).

For the spatial semi discretization in Baker & Dougalis (1980), they trade integrability, coming from the inverse of the discretized differential operator $\Delta_h$, 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 to transfer and extend the results with additional technical effort to the non-autonomous case.

For the full discretization, the proofs in Baker et al. (1979); Baker & Dougalis (1980) rely on an expansion of the discrete error in the basis of $\Delta_h$. 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. Let us also note that the bounds in Baker & Dougalis (1980) derived in the $L^\infty$-norm are of order $k + 1$, but since our bounds involve the $W^{1,\infty}$-norm, we derive optimal error bounds 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 (1975) and Scott (1976). Many extensions and refinements have been achieved in the following years, see, e.g., Nitsche (1975, 1977); Rannacher (1976); Rannacher & Scott (1982); Schatz & Wahlbin (1977, 1982, 1995); Wahlbin (1978).
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 Kashiwabara & Kemmochi (2020). For (evolving) surface finite element methods, estimates on the finite element solution and the generalized Ritz map for isoparametric finite elements are considered in Demlow (2009); Kovács & Power Guerra (2018). In Dörich et al. (2023), the authors of the present paper extended the approach in Brenner & Scott (2008) 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 Bramble et al. (1977); Baker et al. (1977), 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 \( \Delta_b \) on \( L^p \) is shown. This is done either directly using energy techniques, see, e.g., Schatz et al. (1980, 1998), or via resolvent estimates on \( L^p \) and maximal parabolic regularity, see, e.g., Bakaev et al. (2003); Thomée & Wahlbin (2004); Chatzipantelidis et al. (2006); Li (2019). However, such stability estimates cannot be expected for hyperbolic problems in general, see (Arendt et al., 2011, Exa. 8.4.9) and Littman (1963).

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 \( \tau \) such that \( a \leq Cb \). For the sake of readability, we introduce the notation \( t^n = n\tau \) and for an arbitrary time-dependent, continuous object \( x(t) \) in some Banach space \( X \) and a sequence \( (x^n) \) in \( X \), we define

\[ \|x\|_{L^p(X)} := \max_{[0,T]} \|x(t)\|_X , \quad \|x^n\|_{H^r(X)} := \max_{m=1,\ldots,n} \|x^n\|_X . \]

If it is clear from the context, we write \( L^p \) instead of \( L^p(\Omega) \) or \( L^p(\Omega_0) \).

2. General Setting

For a convex, bounded domain \( \Omega \subset \mathbb{R}^N, N = 2,3 \), with boundary \( \partial \Omega \in C^{s,1}, s \in \mathbb{N} \), we study the non-autonomous wave equation (1.1) and the positive, self-adjoint operator \( -\Delta \) on \( L^2(\Omega) \) with homogeneous Dirichlet boundary conditions. Therefore, we introduce the spaces \( H = L^2(\Omega) \) and \( V = \)
\( \mathcal{D}(\mathcal{L}) = H_0^1(\Omega) \). The equation is further equipped with initial values

\[
\begin{align*}
    u(0) &= u^0, \\
    \partial_t u(0) &= v^0.
\end{align*}
\]

We expect that it is also possible to treat more general elliptic differential operators \( \mathcal{L} \) with regular coefficients and Neumann boundary conditions, as long as one can establish the properties of the spatial discretization as stated in Section 2.1. Our analysis relies on the following regularity assumptions of \( \lambda \).

**ASSUMPTION 2.1** There are \( \kappa, \ell_{\max} \in \mathbb{N} \) such that the following holds.

\( (\lambda_1) \) There exist \( C_\lambda \geq c_\lambda > 0 \) such that the function \( \lambda : [0, T] \times \Omega \to \mathbb{R} \) satisfies

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

Moreover, we have \( \lambda, \lambda^{-1} \in C^2([0, T], W^{\kappa, \infty}(\Omega)) \).

\( (\lambda_2) \) For \( 0 \leq \ell \leq \ell_{\max} \) and \( u \in \mathcal{D}(\mathcal{L}^{(\ell)/2}) \) it holds

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

We note that assumption \( (\lambda_2) \) guarantees that the multiplication with \( \lambda \) preserves the boundary conditions incorporated in \( \mathcal{L} \).

**EXAMPLE 2.2** On a smooth domain \( \Omega \) it holds

\[
    \mathcal{D}(\Delta) = \{ u \in H^2(\Omega) \mid u|_{\partial \Omega} = 0 \}, \quad \mathcal{D}(\Delta^2) = \{ u \in H^4(\Omega) \mid u|_{\partial \Omega} = \Delta u|_{\partial \Omega} = 0 \}.
\]

In this case, we have the following sufficient conditions for \( (\lambda_2) \).

(a) We always have \( \ell_{\max} \geq 2 \) and achieve \( \ell_{\max} \geq 4 \) by the product rule if

\[
    \nabla x \lambda|_{\Gamma} = 0. \quad (2.1)
\]

(b) Having the quasilinear case \((1.2)\) in mind, and assuming \( \lambda = \lambda(t, u) \) where \( u \) is the solution in \( H_0^1(\Omega) \), then a sufficient condition for \( (2.1) \) is given by \( \partial_t \lambda(t, 0) = 0 \).

(c) If \( \nabla x \lambda \) has compact support in \( \Omega \), \( (\lambda_2) \) is satisfied for any \( \ell_{\max} \in \mathbb{N} \).

Further, condition \( (\lambda_1) \) 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 \( j = 0, 1, 2 \) the bounds

\[
    \left\| \partial_j^{(\ell)}(t) \phi \right\|_{W^{\ell,p}} \lesssim C \left\| \phi \right\|_{W^{\ell,p}}, \quad \left\| \partial_j^{(\ell)}(t)^{-1} \phi \right\|_{W^{\ell,p}} \lesssim C \left\| \phi \right\|_{W^{\ell,p}},
\]

with a constant \( C > 0 \) depending on \( \lambda \) and its derivatives.

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

\[
    \partial_t y(t) = \Lambda(t)^{-1} A y(t) + F(t), \quad t \in [0, T], \quad (2.2)
\]
with initial value \( y(0) = y^0 \) in the product space \( X = V \times H \), with
\[
 y = \begin{pmatrix} u \\ \partial_t u \end{pmatrix}, \quad y^0 = \begin{pmatrix} u^0 \\ \partial_t u^0 \end{pmatrix}, \quad A(t) = \begin{pmatrix} \text{Id} & 0 \\ 0 & \lambda(t) \end{pmatrix}, \quad A = \begin{pmatrix} 0 & \text{Id} \\ \Delta & 0 \end{pmatrix}, \quad F(t) = \begin{pmatrix} 0 \\ f(t) \end{pmatrix}.
\]

In particular, we emphasize that under Assumption 2.1 the operator \( A \) generates the time-dependent inner product
\[
(y \mid z)_{A(t)} = (A(t)y \mid z)_X, \quad t \in [0,T], y,z \in X.
\]

Since the multiplication with \( \lambda, \lambda^{-1} \) is continuous on \( L^2 \), the corresponding norm is equivalent to the norm of \( X \), i.e., we have
\[
c_A \| z \|_X^2 \leq \| z \|_{A(t)}^2 \leq C_A \| z \|_X^2, \quad t \in [0,T], z \in X,
\]
with constants \( c_A = \min \{ 1, c_\lambda \} \) and \( C_A = \max \{ 1, C_\lambda \} \). Further, we conclude from (\( \lambda_2 \)) the continuity of the map
\[
A(t) : \mathcal{D}(A^\ell) \to \mathcal{D}(A^\ell), \quad 0 \leq \ell \leq \ell_{\text{max}}, t \in [0,T].
\]

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 \( \Delta^{-1} : H \to V \) given by
\[
- (\Delta^{-1} \varphi \mid \psi)_V = (\varphi \mid \psi)_H, \quad \varphi \in H, \psi \in V.
\]

For the analysis, we heavily rely on the following elliptic regularity result (Grisvard, 1985, Thm. 2.4.2.5).

**THEOREM 2.4** (Elliptic regularity) Let \( \partial \Omega \subset C^{1,1} \), then for all \( 1 < p < \infty \) there is a constant \( C_p > 0 \) such that for all \( \varphi \in L^p(\Omega) \) it holds
\[
\| \Delta^{-1} \varphi \|_{W^{2,p}} \leq C_p \| \varphi \|_{L^p}.
\]

Furthermore, we define the first-order solution operator \( T : X \to \mathcal{D}(A) \) by
\[
T = \begin{pmatrix} 0 & \Delta^{-1} \\ \text{Id} & 0 \end{pmatrix}.
\]

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

### 2.1 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 Elliott & Ranner (2021). In particular, we introduce a shape-regular and quasi-uniform mesh \( \mathcal{T}_h \), consisting of isoparametric elements of degree \( k \in \mathbb{N} \) and let \( \partial \mathcal{\Omega} \subset C^{k+1,1} \). The computational domain \( \mathcal{\Omega}_h \) is given by
\[
\mathcal{\Omega}_h = \bigcup_{K \in \mathcal{T}_h} K \approx \mathcal{\Omega},
\]
where the subscript $h$ denotes the maximal diameter of all elements $K \in \mathcal{T}_h$. In the following, we require that $h \leq h_0$ such that all cited results below hold true. We note that $h_0$ only depends on the geometry of the domain $\Omega$ and the polynomial degree $k$. Based on the transformations $F_K$ mapping the reference element $\tilde{K}$ to $K \in \mathcal{T}_h$, we introduce the finite element space of degree $k$

$$W_h = \{ \varphi \in C_0(\mathcal{T}_h) \mid \varphi|_K = \tilde{\varphi} \circ (F_K)^{-1} \text{ with } \tilde{\varphi} \in \mathcal{P}^k(\tilde{K}) \text{ for all } \tilde{K} \in \mathcal{T}_h \}.$$ 

Here, $\mathcal{P}^k(\tilde{K})$ consists of all polynomials on $\tilde{K}$ of degree at most $k$. The discrete approximation spaces are given by

$$H_h = \left( W_h, (\cdot, \cdot)_{L^2(\Omega_h)} \right), \quad V_h = \left( W_h, (\cdot, \cdot)_{H^1(\Omega_h)} \right),$$

and we set $X_h = V_h \times H_h$. Following the detailed construction in (Elliott & Ranner, 2021, Sec. 5), we introduce the lift operator $\mathcal{L}_h : H_h \to H$. In particular, for $p \in [1, \infty]$ there are constants $c_p, C_p > 0$ with

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

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

cf. (Elliott & Ranner, 2021, Prop. 5.8). Further by (Elliott & Ranner, 2013, Sec. 4), the lift preserves node values, i.e., in particular

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

where we denote the nodal interpolation operator by $I_h : C(\Omega) \to V_h$. As shown in (Elliott & Ranner, 2021, Thm. 5.9), we have for $m = 0, 1, 1 \leq p \leq \infty$, and $1 \leq \ell \leq k$ the estimates

$$\| (I - \mathcal{L}_h I_h) \varphi \|_{W^{m,p}(\Omega)} \lesssim h^{\ell+1-m} \| \varphi \|_{W^{\ell+1,p}(\Omega)}, \quad \varphi \in W^{\ell+1,p}(\Omega). \quad (2.8)$$

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

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

$$\begin{align*}
(\mathcal{L}_h^H \varphi | \psi_h)_{H_h} &= (\varphi | \mathcal{L}_h \psi_h)_H, \quad \varphi \in H, \psi_h \in H_h, \quad (2.9a) \\
(\mathcal{L}_h^V \varphi | \psi_h)_{V_h} &= (\varphi | \mathcal{L}_h \psi_h)_V, \quad \varphi \in V, \psi_h \in V_h, \quad (2.9b)
\end{align*}$$

and note in the conforming case $\mathcal{L}_h^H$ and $\mathcal{L}_h^V$ coincide with the $L^2$- and the Ritz projection, respectively. From (Hipp et al., 2019, Thm. 5.3) and (Elliott & Ranner, 2021, Lem. 8.24), we obtain for $1 \leq \ell \leq k$

$$\| \mathcal{L}_h^H \varphi \|_{H_h} \lesssim \| \varphi \|_{L^2(\Omega)}, \quad \varphi \in L^2(\Omega), \quad (2.10a)$$

$$\| (I_h - \mathcal{L}_h^H) \varphi \|_{H_h} \lesssim h^{\ell+1} \| \varphi \|_{H^{\ell+1}(\Omega)}, \quad \varphi \in H^{\ell+1}(\Omega), \quad (2.10b)$$

as well as for $1 \leq \ell \leq k$

$$\| \mathcal{L}_h^V \varphi \|_{V_h} \lesssim \| \varphi \|_{H^1(\Omega)}, \quad \varphi \in H^1(\Omega), \quad (2.11a)$$

$$\| (I_h - \mathcal{L}_h^V) \varphi \|_{H_h} \lesssim h^{\ell+1} \| \varphi \|_{H^{\ell+1}(\Omega)}, \quad \varphi \in H^{\ell+1}(\Omega). \quad (2.11b)$$

In addition, we need the stability of $\mathcal{L}_h^H$ in $H^1$

$$\| \mathcal{L}_h^H \varphi \|_{V_h} \lesssim \| \varphi \|_{H^1(\Omega)}, \quad \varphi \in H^1(\Omega), \quad (2.12)$$
Further, we define the (standard) $L^2$-projection $\pi_h$ onto $V_h$ for $\varphi \in L^2(\Omega_h)$ via

$$\langle \pi_h \varphi \mid \psi_h \rangle_{L^2(\Omega_h)} = \langle \varphi \mid \psi_h \rangle_{L^2(\Omega_h)}, \quad \varphi \in L^2(\Omega_h), \psi_h \in V_h.$$ 

We note that it only differs from $L^2(\Omega)$ by geometric errors and thus, we obtain for $\varphi \in H^2(\Omega)$

$$\|\pi_h \varphi - \varphi\|_{L^2(\Omega_h)} \lesssim h^{k+1} \|\varphi\|_{H^2(\Omega)}.$$ \hfill (2.13)

The proofs for (2.12) and (2.13) are given in Appendix A. In addition, there hold the following stability estimates. Let $p \in [2, \infty]$, then for any $\varphi \in L^p(\Omega_h)$ and $\psi \in H^1(\Omega_h)$

$$\|\pi_h \varphi\|_{L^p(\Omega_h)} \leq C \|\varphi\|_{L^p(\Omega_h)}, \quad \|\pi_h \psi\|_{H^1(\Omega_h)} \leq C \|\psi\|_{H^1(\Omega_h)},$$ \hfill (2.14)

with a constant $C > 0$ independent of $h$. We note that the case $p = 2$ is trivially satisfied by the definition of the projection. The case $p = \infty$ is covered by (Nitsche, 1975, Thm. 1), and an interpolation argument yields the first bound. The stability in $H^1$ is shown as for (2.12).

For the analysis in the following sections, we additionally rely on stability and approximation properties of $L^2(\Omega)$ in the maximum norm. These features are well known in the literature for conforming finite elements, see, e.g., the monograph (Brenner & Scott, 2008, Ch. 8). In the non-conforming case, the authors recently extended these results in (Dörich et al., 2023, Thm. 2.5 & 2.6) to isoparametric finite elements. A special case is stated in the following proposition.

**Proposition 2.5** Let $\partial \Omega \in C^{k+1,1}$ and $h \leq h_0$. Then, the adjoint lift is stable in $W^{1,\infty}$ with

$$\|L^2_h \varphi\|_{W^{1,\infty}(\Omega_h)} \lesssim \|\varphi\|_{W^{1,\infty}(\Omega)}, \quad \varphi \in W^{1,\infty}(\Omega).$$ \hfill (2.15)

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

$$\|\text{Id} - L^2_h L^\ell_h \varphi\|_{W^{1,\ell}(\Omega_h)} \lesssim h^\ell \|\varphi\|_{W^{\ell+1,\ell}(\Omega)}, \quad \varphi \in W^{\ell+1,\ell}(\Omega).$$ \hfill (2.16)

Related estimates can be found in Kashiwabara & Kemmochi (2020) for the Neumann problem and linear elements, and in the context of evolving surfaces also in Kovács & Power Guerra (2018). We will also employ the inverse estimate, cf. (Brenner & Scott, 2008, Thm. 4.5.11) or (Maier, 2020, Lem. 5.6)

$$\|\varphi_h\|_{L^\infty(\Omega_h)} \leq C h^{-\frac{k-1}{p}} \|\varphi_h\|_{L^p(\Omega_h)}, \quad \|\varphi_h\|_{V_h} \leq C h^{-1} \|\varphi_h\|_{L^2(\Omega_h)}, \quad \varphi_h \in V_h.$$ \hfill (2.17)

We introduce the first-order lift operator $L^\ell_h: W^{\ell+1,\ell}(\Omega_h)^2 \rightarrow W^{\ell+1,\ell}(\Omega)^2$, for $\ell = 0, 1$ and $1 \leq p \leq \infty$, and reference operator $J_h: V \times V \rightarrow X_h$ defined by

$$L^\ell_h = \begin{pmatrix} L^\ell_h & 0 \\ 0 & L^\ell_h \end{pmatrix}, \quad J_h = \begin{pmatrix} L^\ell_h & 0 \\ 0 & L^\ell_h \end{pmatrix},$$

which are bounded uniformly in $h$ due to (2.7) and (2.11). In particular, we have by (2.16) the stability $J_h \in L((W^{1,\infty}(\Omega))^2, (W^{1,\infty}(\Omega_h))^2)$. For $t \in [0, T]$ we define the discrete operators $\lambda_h(t): H_h \rightarrow H_h$, $\Lambda_h(t): X_h \rightarrow X_h$, and the discrete right-hand side $F_h(t)$ by

$$\lambda_h(t) \varphi_h = \pi_h(I_h \lambda(t) \varphi_h), \quad \Lambda_h(t) = \begin{pmatrix} I_h(t) & 0 \\ 0 & \lambda_h(t) \end{pmatrix}, \quad F_h(t) = \begin{pmatrix} 0 \\ I_h(t) f(t) \end{pmatrix}.$$ 

Correspondingly to Lemma 2.3, we collect important properties of $\lambda_h$ in the following lemma.
LEMMA 2.6 Let Assumption 2.1 be satisfied for some $\kappa \geq 2$ and $h \leq h_0$. Then, we have for $t \in [0, T]$, $1 \leq p \leq \infty$, and $j = 0, 1, 2$ the bounds
\[
\|\partial_t^j \lambda_h(t) \phi_h\|_{L^p} \leq C \|\phi_h\|_{L^p}, \quad \|\partial_t^j \lambda_h(t) \phi_h\|_{V_h} \leq C \|\phi_h\|_{V_h},
\]
with a constant $C > 0$ depending only on $\lambda$ and its derivatives.

Proof. Using the stability in (2.14), it is sufficient to show the assertion for the product $(I_h \partial_t^j \lambda) \phi_h$. Combining the interpolation property (2.8) and Assumption 2.1, the Hölder inequality yields the desired bound.

Finally, we introduce the operators $\Delta_h : V_h \to H_h$ and $A_h : X_h \to X_h$ for $\phi_h, \psi_h \in V_h$ given by
\[
- (\Delta_h \phi_h \mid \psi_h)_{H_h} = (\phi_h \mid \psi_h)_{V_h}, \quad \quad A_h = \begin{pmatrix} 0 & \text{Id} \\ \Delta_h & 0 \end{pmatrix}.
\]
Note that these operators are not uniformly bounded with respect to $h$. Correspondingly to (2.3) and (2.4) using the identity $(\lambda_h \phi_h \mid \psi_h)_{L^2(\Omega_h)} = (I_h \lambda \phi_h \mid \psi_h)_{L^2(\Omega_h)}$, the discrete operator $A_h$ generates the time-dependent inner product
\[
(\gamma_h \mid z_h)_{\Delta_h(t)} = (A_h(t) \gamma_h \mid z_h)_{X_h}, \quad t \in [0, T], \gamma_h, z_h \in X_h,
\]
with the induced norm being equivalent to the norm of $X_h$, i.e., we have as in (2.4) by Lemma 2.6
\[
c_{A_h} \|z_h\|_{X_h}^2 \leq \|z_h\|_{\Delta_h(t)}^2 \leq C_{A_h} \|z_h\|_{X_h}^2, \quad t \in [0, T], z_h \in X_h.
\] (2.18)

We define the discrete solution operator $\Delta_h^{-1} : H_h \to V_h$ by
\[
- (\Delta_h^{-1} \phi_h \mid \psi_h)_{V_h} = (\phi_h \mid \psi_h)_{H_h}, \quad \phi_h, \psi_h \in V_h,
\] (2.19)
and further the corresponding first-order solution operator
\[
T_h = \begin{pmatrix} 0 & \Delta_h^{-1} \\ \text{Id} & 0 \end{pmatrix},
\]
which again satisfies $T_h A_h = \text{Id}$ and $A_h T_h = \text{Id}$ on $X_h$.

The spatially discrete non-autonomous wave equation in first-order formulation then reads
\[
\partial_t y_h(t) = A_h(t)^{-1} A_h y_h(t) + F_h(t), \quad t \in [0, T],
\] (2.20)
with the initial value $y_h(0) = y^0_h$, where
\[
y^0_h = J_h y^0 = (\mathcal{L}_h^{\text{V}^*} u^0, \mathcal{L}_h^{\text{V}^*} v^0)^T.
\] (2.21)

This choice of the initial value guarantees the convergence of the expression $J_h A_h^2 y^0 - A_h^2 J_h y^0$ with optimal order, see Lemma 5.4 below, and error bounds for the first approximations, see Section 5.3. We note that in practice, one would usually choose the interpolation $y_h(0) = (I_h u^0, I_h v^0)$. However, already in Baker & Dougalis (1980) preconditioned initial values of the type $u^0_h = (\Delta_h^{-1})^m \Delta^m u^0$, $m \geq 4$, had to be chosen for the full order of convergence. Note that (2.21) is equivalent to the case $m = 1$ in
the conforming case, but we do not require knowledge on $\Delta u^0$. A discussion on the computation of high-order approximations to $\mathcal{L}_h^{V^s}$ is given in (Dörich, 2022, Prop. 2.4).

In the spatially continuous case, the solution operator $\Delta^{-1}$ can be used to obtain regularity which is traded in for pointwise estimates via the bounded map

$$\Delta^{-1} : L^2 \rightarrow H^2 \rightarrow L^\infty.$$ 

However, since we use Lagrangian finite elements which are not $H^2$-conforming, this approach does not work with $\Delta_h^{-1}$. Hence, in the following we provide estimates of $\Delta_h^{-1}$ that directly give us integrability without a detour via higher-order Sobolev spaces. A weaker form of this result has already been proven in (Bramble et al., 1977, Lem. 4.1) in the conforming case only, and was recently sharpened in Dörich (2022). We state a variant of this result and, for completeness, give the proof in Appendix A.

**Lemma 2.7** Let $\partial \Omega \in C^{1,1}$ and $h \leq h_0$. Then, the solution operator $\Delta_h^{-1}$ satisfies

$$\| \Delta_h^{-1} \varphi_h \|_{L^\infty} \lesssim \| \varphi_h \|_{L^2} \quad \text{and} \quad \| \Delta_h^{-1} \varphi_h \|_{W^{1,\infty}} \lesssim \| \varphi_h \|_{L^4}$$

for $\varphi_h \in V_h$.

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

$$X_h \hookrightarrow L^4 \times L^2 \xrightarrow{T_h} L^\infty \times L^4 \xrightarrow{T_h} W^{1,\infty} \times L^\infty,$$

which allow us to bound the maximum norm $\| \cdot \|_{W^{1,\infty} \times L^\infty}$ in terms of the energy norm $\| \cdot \|_X$ if we apply the solution operator $T_h$ twice. We explain in Section 3 how to employ this observation. Note that in fact, one could sharpen the result to show that $\Delta_h^{-1} : L^2 \rightarrow W^{1,p}$ is bounded uniform in $h$ as long as the embedding $H^2 \rightarrow W^{1,p}$ is valid. For $p > N$, this then also implies the first estimate of the lemma.

We can finally state our first main result on the semi discretization. The proof is given in Section 3. We let $\mathcal{H}_A^k := \mathcal{D}\left((-\Delta)^{k/2}\right)$ and use the notation

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

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

**Theorem 2.8** Let $\partial \Omega \in C^{k+1,1}$, $h \leq h_0$, and let Assumption 2.1 hold for some $\ell_{\max} \in \mathbb{N}$ and $\kappa = k + 1$. Further, assume that the right-hand side $f$ and the solution of (1.1) satisfy

$$u \in C^4([0,T], H^{k'}(\Omega)) \cap C^2([0,T], H^{k+1}(\Omega)) \cap C^1([0,T], W^{k+1,\infty}(\Omega)),$$

$$f \in C^1([0,T], H^{k+1}(\Omega)) \cap C([0,T], \mathcal{H}_A^k),$$

$$u^0 \in H^{k+2}(\Omega) \cap \mathcal{H}_A^3, \quad v^0 \in H^{k+1}(\Omega) \cap \mathcal{H}_A^2,$$

and the initial value $y_h(0)$ is chosen as in (2.21). Then we have the error bound

$$\| y(t) - \mathcal{L}_h y_h(t) \|_{W^{1,\infty} \times L^\infty} \leq C h^k,$$

where $C$ is independent of $h$. 

2.2 Full discretization

We study the full discretization with the backward Euler scheme

$$\partial_t y^n_h = A_h^{-1}(t^n) A_h y^n_h + F_h(t^n), \quad n \geq 1,$$

(2.24)

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

$$\partial_t \phi^n = \frac{\phi^n - \phi^{n-1}}{\tau}.
$$

(2.25)

For the fully discrete scheme, we use the same initial value and set $y_h^0 = J_h y^0 = (L_h^{V^*} u^0, L_h^{V^*} v^0)^T$ as in (2.21). Our second main result on the full discretization, which is proved in Section 4, then reads as follows.

**Theorem 2.9** Let $\partial \Omega \in C^{k+1,1}$, $h \leq h_0$, and let Assumption 2.1 hold for $\ell_{\text{max}} \geq 2$ and $\kappa = k + 1$. Further, assume that the right-hand side $f$ and the solution of (1.1) satisfy in addition to (2.23)

$$u \in C^5([0, T], H^1(\Omega)) \cap C^4([0, T], H^2_A) \cap C^3([0, T], \mathcal{H}^3_A),$$

$$f \in C^1([0, T], \mathcal{H}^k_A) \cap C([0, T], \mathcal{H}^{k+1}_A),$$

$$u^0 \in \mathcal{H}^{k+3}_A, \quad v^0 \in \mathcal{H}^{k+2}_A,$$

and the initial value $y_h^0$ is chosen as in (2.21). Then, there is $\tau_0 > 0$ such that for $\tau \leq \tau_0$ we have the error bound

$$\|y(t^n) - \mathcal{L}_h y_h^n\|_{W^{1,-\infty}L^\infty} \leq C \tau + Ch_{\text{min}}^{\min\{k, \ell_{\text{max}}\}}, \quad n \geq 2,$$

where $C$ is independent of $h$ and $\tau$, and $\tau_0$ is independent of $h$.

We refer to Remark 5.10 below in order to explain the minimum in the convergence rate. Further, we emphasize that the first two approximations do not enter the above error bound. However, we have the following convergence result for the first approximations.

**Theorem 2.10** Let the assumptions of Theorem 2.9 hold. Then, we have

$$\|y(t^\ell) - \mathcal{L}_h y_h^\ell\|_{W^{1,-\infty}L^\infty} \leq C (\tau + h^k), \quad \ell = 0, 1.$$

The proof is given in Section 5.

**Remark 2.11** Considering the strategy of the proof, we see that the same ideas can be applied by only differentiating the error equations once and to exploit the relation

$$X_h = H^1 \times L^2 \overset{T_h}{\longrightarrow} L^\infty \times H^1.$$

Following the lines of the presented proof below, we obtain the very same estimates for the $L^\infty \times H^1$ norm. The gain in using the "weaker" norm are the decreased regularity assumptions as well as the simpler choice of the initial value $y_h^0 = (L_h^{V^*} u^0, L_h v^0)^T$, where we save one computation of $\mathcal{L}_h^{V^*}$. 
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 (2.20) in the maximum norm. To this end, we proceed as follows. We split the error in
\[ y(t) - \mathcal{L}_h y_h(t) = (\mathbb{I} - \mathcal{L}_h J_h) y(t) + \mathcal{L}_h (J_h y(t) - y_h(t)) =: e_h(t) + \mathcal{L}_h e_h(t) \tag{3.1} \]
and derive an equation for the discrete error $e_h$. With the solution operator $T$, we rewrite (2.2) as
\[ T \Lambda(t) \partial_y y = y + T \Lambda(t) F(t), \quad t \in [0, T], \]
with initial value $y(0) = y^0$. Correspondingly, we use the discrete solution operator $T_h$ to obtain from (2.20) the semi-discrete equation
\[ T_h A_h(t) \partial_y y_h = y_h + T_h A_h(t) F_h(t), \quad t \in [0, T], \tag{3.2} \]
with initial value $y_h(0) = y^0_h$. Thus, we conclude that the discrete error $e_h$ solves the evolution equation
\[ T_h A_h(t) \partial_y e_h(t) = e_h(t) + \delta_h(t), \quad t \in [0, T], \tag{3.3} \]
with initial value $e_h(0) = e^0_h = J_h y_0 - y^0_h$ and the defect
\[ \delta_h(t) = (T_h A_h(t) - J_h T \Lambda(t)) \partial_y y(t) + J_h y(t) \Lambda(t) F(t) - T_h A_h(t) F_h(t). \tag{3.4} \]

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 (2.22). Note that we can view the error equation as an ordinary differential equation in a finite-dimensional space with right-hand side, say $g(t, e_h)$. Since the right-hand side $g$ is smooth by our assumptions, we obtain local existence. Since all norms are equivalent, the error bounds below guarantee existence up to the final time $T$.

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 choice (2.21) of the discrete initial value $y^0_h$. Here, we obtain from (2.2) and (2.20) for the discrete error $e_h$ the evolution equation
\[ A_h(t) \partial_y e_h(t) = A_h e_h(t) + \delta_{h,A}(t), \quad t \in [0, T], \tag{3.5} \]
with the defect
\[ \delta_{h,A}(t) = (A_h(t) - J_h A(t)) \partial_y y(t) + (J_h A - A_h J_h) y(t) + J_h A(t) F(t) - A_h(t) F_h(t). \tag{3.6} \]
Note that we have the relation $\delta_{h,A} = A_h \delta_{h,T}$. Moreover, we emphasize that a similar defect was already studied in the unified error analysis provided in Hipp et al. (2019). However, here we also have to bound time derivatives of $\delta_{h,T}$ and $\delta_{h,A}$ as well as a different choice of $J_h$. We postpone the derivation of these bounds as well as the estimates for the time derivatives of the initial error to Section 5.
Finally, the Gronwall inequality implies for all $j$

\[
\|e_h(t)\|_{W^{1,r} \times L^r} \lesssim \left\| \frac{\partial^j e_h}{\partial t^j} \right\|_{X_0} + \left\| \frac{\partial^{j+1} e_h}{\partial t^{j+1}} \right\|_{X_0} + \left\| \frac{\partial^j \delta_{h,T}}{\partial t^j} \right\|_{X_0} + \left\| \frac{\partial^j \delta_{h,A}}{\partial t^j} \right\|_{X_0}.
\]

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

\[
\|e_h\|_{L_t^r(W^{1,r} \times L^r)} \lesssim \left\| \frac{\partial^1 e_h}{\partial t^1} \right\|_{L_t^r(W^{1,r} \times L^r)} + \left\| \frac{\partial^2 e_h}{\partial t^2} \right\|_{L_t^r(W^{1,r} \times L^r)} + \left\| \frac{\partial^j \delta_{h,T}}{\partial t^j} \right\|_{L_t^r(W^{1,r} \times L^r)} + \left\| \frac{\partial^j \delta_{h,A}}{\partial t^j} \right\|_{L_t^r(W^{1,r} \times L^r)}.
\]

**Proof.** From the representation in (3.3) we obtain with the properties of $T_h$ in Lemma 2.7 and (2.22), and the bounds on $A_h$ in Lemma 2.6

\[
\|e_h(t)\|_{W^{1,r} \times L^r} \lesssim \left\| \frac{\partial^1 e_h}{\partial t^1} \right\|_{L_t^r(W^{1,r} \times L^r)} + \left\| \frac{\partial^j \delta_{h,T}}{\partial t^j} \right\|_{L_t^r(W^{1,r} \times L^r)}.
\]

For the derivative we compute using the same estimates

\[
\left\| \frac{\partial^1 e_h}{\partial t^1} \right\|_{L_t^r(W^{1,r} \times L^r)} \lesssim \left\| T_h A_h(t) \frac{\partial^1 e_h}{\partial t^1} \right\|_{L_t^r(W^{1,r} \times L^r)} + \left\| T_h \frac{\partial^1 A_h}{\partial t} e_h(t) \right\|_{L_t^r(W^{1,r} \times L^r)} + \left\| \frac{\partial^1 \delta_{h,T}}{\partial t} \right\|_{L_t^r(W^{1,r} \times L^r)} \lesssim \left\| \frac{\partial^2 e_h}{\partial t^2} \right\|_{X_0} + \left\| \frac{\partial^2 A_h}{\partial t} e_h(t) \right\|_{X_0} + \left\| \frac{\partial^1 \delta_{h,A}}{\partial t} \right\|_{X_0},
\]

and, using Sobolev’s embedding $\|\xi_h\|_{L_t^4 W_x^1} \lesssim \|\xi_h\|_{X_0}$ for $\xi_h \in X_0$, concludes the proof. \(\square\)

In the following lemma, we provide the bounds of the time derivatives appearing in Lemma 3.1 in the $X_0$ 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 $j = 1, 2$ we have

\[
\left\| \frac{\partial^j e_h}{\partial t^j} \right\|_{L_t^r(W^{1,r} \times L^r)} \leq e^{CT} \sum_{i=1}^{j} \left( \left\| \frac{\partial^i e_h(0)}{\partial t^i} \right\|_{X_0} + \left\| \frac{\partial^i \delta_{h,A}}{\partial t^i} \right\|_{L_t^r(W^{1,r} \times L^r)} \right).
\]

**Proof.** In the following, we prove for $j = 1, 2$ the estimate

\[
\left\| \frac{\partial^j e_h}{\partial t^j} \right\|_{L_t^r(W^{1,r} \times L^r)} \leq (1 + T) e^{CT} \left( \left\| \frac{\partial^j e_h(0)}{\partial t^j} \right\|_{X_0} + \left\| \frac{\partial^j \delta_{h,A}}{\partial t^j} \right\|_{L_t^r(W^{1,r} \times L^r)} + \sum_{i=1}^{j-1} \left\| \frac{\partial^i e_h}{\partial t^i} \right\|_{L_t^r(W^{1,r} \times L^r)} \right).
\]

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

\[
\sum_{i=0}^{j-1} \left( \int_{t_i}^{t_{i+1}} \frac{J}{J} A_h \frac{\partial}{\partial t} e_h - A_h \frac{\partial}{\partial t} e_h = \frac{\partial}{\partial t} \delta_{h,A},
\right)
\]
for $j = 0, 1, 2$. Taking the inner product with $\partial_t^j e_h$ gives

$$\left( A_h \partial_t^{j+1} e_h \mid \partial_t^j e_h \right)_{X_h} = \left( A_h \partial_t^j e_h \mid \partial_t^j e_h \right)_{X_h} + \left( \partial_t^j \delta_{h,A} \mid \partial_t^j e_h \right)_{X_h} - \sum_{\ell=0}^{j-1} \left( \partial_t^{j-\ell} A_h \partial_t^{\ell+1} e_h \mid \partial_t^j e_h \right)_{X_h}. $$

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

$$2 \left( A_h \partial_t^{j+1} e_h \mid \partial_t^j e_h \right)_{X_h} \leq \| \partial_t^j \delta_{h,A} \|_{X_h}^2 + 2 \| \partial_t^j e_h \|_{X_h}^2 + \sum_{\ell=0}^{j-1} \left( \partial_t^{j-\ell} A_h \partial_t^{\ell+1} e_h \|_{X_h}^2. $$

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

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

with the constants

$$C_j = 2^j + j \| \partial_t A_h \|_{L^\infty(L^2(X_h))}, \quad \widehat{C}_j = \max_{\ell=0, \ldots, j-1} \left( \frac{j}{\ell} \| \partial_t^{j-\ell} A_h \|_{L^\infty(L^2(X_h))}. $$

Note that these constants are bounded independently of $j \leq 2$ by $C_2$ and $\widehat{C}_2$, respectively.

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

$$\frac{d}{dt} \| \partial_t^j e_h \|_{A_h(t)}^2 = 2 \left( A_h \partial_t^{j+1} e_h \mid \partial_t^j e_h \right)_{X_h} + \left( \partial_t A_h \partial_t^j e_h \mid \partial_t^j e_h \right)_{X_h}. $$

Moreover, integration in time, using the boundedness of $\partial_t A_h$ for the second term, and the norm equivalence (2.18) yields

$$\| \partial_t^j e_h \|_{A_h(t)}^2 \leq \| \partial_t^j e_h(0) \|_{A_h(0)}^2 + t \| \partial_t^j \delta_{h,A} \|_{L^\infty(X_h)}^2 + t \sum_{\ell=1}^{j-1} \| \partial_t^\ell e_h \|_{L^\infty(X_h)}^2 + \int_0^t \| \partial_t^j e_h(t) \|_{L^\infty(A_h(t))} ds. $$

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

$$\| \partial_t^j e_h(t) \|_{A_h(t)}^2 \leq e^{C_j} \left( \| \partial_t^j e_h(0) \|_{A_h(0)}^2 + t \| \partial_t^j \delta_{h,A} \|_{L^\infty(X_h)}^2 + \sum_{\ell=1}^{j-1} \| \partial_t^\ell e_h \|_{L^\infty(X_h)}^2 \right),$$

and (3.8) follows with (2.18).

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 property derived in (2.16)

$$\| y(t) - \mathcal{L}_h y_h(t) \|_{W^{1,\infty} L^\infty} \leq \| e_h(t) \|_{W^{1,\infty} L^\infty} + C_{\mathcal{L}_h} \| e_h(t) \|_{W^{1,\infty} L^\infty} \leq C h^k \| y(t) \|_{W^{k+1,\infty} W^{1,\infty} L^\infty} + C_{\mathcal{L}_h} \| e_h(t) \|_{W^{1,\infty} L^\infty},$$

and apply Lemmas 3.1 and 3.2. The remaining defects and errors in the initial values are bounded in Lemmas 5.2, 5.5, and 5.6. \qed
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.

4.1 Calculus for discrete derivatives

We first need some auxiliary results for the discrete derivatives defined in (2.25). A straightforward calculation yields the following.

**Lemma 4.1** It holds the discrete product rule

\[ \partial \tau (\varphi^n \psi^n) = (\partial \tau \varphi^n) \psi^n + \varphi^{n-1} (\partial \tau \psi^n) \]

and also the more general discrete Leibniz rule

\[ \partial \tau (\varphi^n \psi^n) = \sum_{\ell=0}^{j} \binom{j}{\ell} (\partial \tau^{j-\ell} \varphi^{n-\ell}) (\partial \tau^{\ell} \psi^n), \quad j \geq 0. \]

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 \((\varphi^n)\) be a sequence in a Hilbert space with inner product \((\cdot | \cdot)\), and let \(k_0 \in \mathbb{N}\).

(a) For any \(M \geq k_0\), it holds

\[ \frac{1}{2} \| \varphi^M \|^2 \leq \frac{1}{2} \| \varphi^{k_0-1} \|^2 + \tau \sum_{j=k_0}^{M} (\partial \tau \varphi^j | \varphi^j). \]

(b) If there are constants \(\alpha, \beta_1, \beta_2 \geq 0\) such that

\[ (\partial \tau \varphi^j | \varphi^j) \leq \alpha^2 + \beta_1 \| \varphi^{j-1} \|^2 + \beta_2 \| \varphi^j \|^2, \quad j \geq k_0, \]

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

\[ \| \varphi^M \| \leq \left( \sqrt{1 + 2\tau \beta_1} \| \varphi^{k_0-1} \| + \sqrt{2\tau \alpha} \right) e^{2(\beta_1 + \beta_2)\tau \alpha}. \]
Proof. Part (a) is for example shown in (Hochbruck & Pažur, 2017, Lemma 4.2). Inserting (4.1) in (a) yields
\[ \| \phi_N \|^2 \leq \| \phi_{k_0} \|^2 + 2 \tau \sum_{j=k_0}^N \left( \alpha^2 + \beta_1 \| \phi_{j-1} \|^2 + \beta_2 \| \phi_j \|^2 \right) \]
\[ \leq (1 + 2 \tau \beta_1) \| \phi_{k_0} \|^2 + 2t_N \alpha^2 + 2(\beta_1 + \beta_2) \tau \sum_{j=k_0}^N \| \phi_j \|^2 \]
and by a Gronwall argument, see, e.g., (Linz, 1969, Lemma 1), we obtain
\[ \| \phi_N \|^2 \leq \left( (1 + 2 \tau \beta_1) \| \phi_{k_0} \|^2 + 2t_N \alpha^2 \right) e^{(\beta_1 + \beta_2)\tau N}. \]
Taking roots yields the assertion. \qed

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 \( x : [0,T] \rightarrow Z \) be \( j \)-times differentiable with bounded derivatives, then
\[ \left\| \partial_j^\ell x(t^n) \right\|_Z \leq \sup_{t \in [t^{n-1},t^n]} \left\| \partial_j^\ell x(t) \right\|_Z. \]
Proof. This simply follows from an iterative application of the fundamental theorem of calculus. \qed

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)
\[ T_h A_h(t^n) J_h \partial y(t^n) = J_h \lambda y(t^n) + T_h A_h(t^n) F_h(t^n) + \delta_{h,T}^n \]
with a defect of the form, using the representation of \( \delta_{h,T} \) in (3.4),
\[ \delta_{h,T}^n = \delta_{h,T}(t^n) + T_h A_h(t^n) J_h \left( \partial y(t^n) - \partial y(t^n) \right). \]  
(4.2)
From this we obtain the fully discrete error equation
\[ T_h A_h(t^n) \partial e_h^n = e_h^n + \delta_{h,T}^n. \]  
(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 (2.24) and obtain
\[ A_h(t^n) J_h \partial y(t^n) = A_h J_h y(t^n) + F_h(t^n) + \delta_{h,A}^n \]
with the defect, using the representation of \( \delta_{h,A} \) in (3.6),
\[ \delta_{h,A}^n = \delta_{h,A}(t^n) + A_h(t^n) J_h \left( \partial y(t^n) - \partial y(t^n) \right). \]  
(4.4)
This gives us the second version of the error recursion
\[ A_h(t^n) \partial e_h^n = A_h e_h^n + \delta_{h,A}^n. \]  
(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 $h$, $\tau$, and $n$ such that
\[
\|e^n_h\|_{W^{1,-\infty}} \lesssim \|\partial^1_t e^n_h\|_{X_h} + \|\partial^2_t e^n_h\|_{X_h} + \|\delta^n_{h,T}\|_{W^{1,-\infty}} + \|\partial^1_T \delta^n_{h,T}\|_{L^{\infty}}
\]
holds for $n \geq 2$.

REMARK 4.5 From this lemma it becomes clear that this technique does not provide bounds for $n = 0, 1$, since we can evaluate $\partial^2_T e^n_h$ only for $n \geq j$.

**Proof of Lemma 4.4.** As in (3.7), we obtain from (4.3) and Lemma 4.1
\[
\partial^1_t e^n_h = -\partial^1_T \delta^n_{h,T} + T_h A_h (t^n - 1) \partial^2_T e^n_h + T_h \partial^1_t A_h (t^n) \partial^1_T e^n_h,
\]
and the proof follows along the lines of Lemma 3.1. \qed

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 and defects.

LEMMA 4.6 Let the assumptions of Theorem 2.9 hold. Then, there is $\tau_0 > 0$ such that for $\tau \leq \tau_0$
\[
\|\partial^j_t e^n_h\|_{X_h} \leq C (1 + T) \sum_{\ell = 1}^j \left( \|\partial^\ell_T e^n_h\|_{X_h} + \|\partial^\ell_T \delta^n_{h,T}\|_{L^{\infty}(X_h)} \right),
\]
for $j = 1, 2$ and $n \geq j + 1$.

**Proof.** As in the proof of Lemma 3.2, we provide the bound
\[
\|\partial^j_t e^n_h\|_{X_h} \leq C (1 + T) \sum_{\ell = 1}^j \left( \|\partial^\ell_T e^n_h\|_{X_h} + \|\partial^\ell_T \delta^n_{h,T}\|_{L^{\infty}(X_h)} + \sum_{\ell = 1}^{j-1} \|\partial^\ell_T e^n_h\|_{X_h} \right), \quad (4.6)
\]
cf. (3.8). Using this estimate recursively directly yields the assertion.

To do so, we apply the bounds from Lemma 4.2 for $\varphi^n = A_h^{1/2} (t^n) \partial^j_t e^n_h$. In particular, we first study the term
\[
(\partial^j_T \varphi^n, e^n_h)_{X_h} = \left( A_h (t^n) \partial^j_T \delta^n_{h,T}, e^n_h \right)_{X_h} + \left( (\partial^j_T A_h^{1/2} (t^n)) \varphi^n, \varphi^n \right)_{X_h},
\]
where we used Lemma 4.1 and the fact that $A_h^{1/2}$ is self-adjoint in $X_h$. Due to Assumption 2.1 and Young’s inequality, this implies
\[
(\partial^j_T \varphi^n, e^n_h)_{X_h} \leq \left( A_h (t^n) \partial^j_T \delta^n_{h,T}, e^n_h \right)_{X_h} + C \|\varphi^n\|_{X_h}^2 + C \|\varphi^n\|_{X_h}^2. \quad (4.7)
\]
For the first term, we obtain with Lemma 4.1 applied to (4.5)
\[
A_h (t^n) \partial^j_T \delta^n_{h,T} = A_h \partial^j_T e^n_h + \partial^j_T \delta^n_{h,T} - \sum_{\ell = 0}^{j-1} \left( \frac{j}{\ell} \right) (\partial^j_{t-\ell} A_h (t^n-\ell)) (\partial^\ell_T e^n_h).
\]

Thus, taking the inner product in $X_h$ with $\partial^j_T e^n_h$ yields as in (3.9) the estimate
\[
2 \left( A_h (t^n) \partial^j_T \delta^n_{h,T}, e^n_h \right)_{X_h} \leq \|\partial^j_T \delta^n_{h,T}\|_{X_h}^2 + C \|\varphi^n\|_{X_h}^2 + C \|\varphi^n\|_{X_h}^2,
\]
Lemma 2.3 where we used Lemma 2.7 together with (5.1a) and (5.1c). Differentiating and taking norms, gives with 
and first consider

Proof. Let the solution

and in the same way for

and the bound (5.1a) and (5.1b) follow from (2.10) and (2.11). For the bound on (5.1c) we exploit the 

18 of 31

B. Dörich, J. Leibold, and B. Maier

5.2 Let the solution

(a) The defect

δ

satisfies for

h

mas 3.1 and 4.4.

Similarly, (5.2b) follows from

and apply Lemmas 4.4 and 4.6 for

n \geq 2. The remaining defects and errors in the initial values are bounded in Lemmas 5.2, 5.5, and 5.7.

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 this section, we mostly omit the time dependency for the sake of readability and assume

h \leq h_0. Further, we take the assumptions of Section 2 as given, but will be precise about the regularity of the solution

u

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. We recall the notation

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

Lemma 5.1 Let

ξ = (φ, ψ) \in H^1(Ω) \times H^{k+1}(Ω)

and λ, f \in C^1([0, T], H^{k+1}(Ω)). Then, the discrete operators introduced in Section 2 satisfy the bounds

\| (T_h J_h - J_h T) ξ \|_{L^∞ L^2} \lesssim h^k \| ξ \|_{H^{k+1}}, \quad (5.1a)

\| (T_h J_h - J_h T) ξ \|_{L^∞ L^2} \lesssim h^k \| ξ \|_{H^{k+1}}, \quad (5.1b)

\| (A_h J_h - J_h A) ξ \|_{L^∞ L^2} \lesssim h^k \| ξ \|_{H^{k+1}}. \quad (5.1c)

For

j = 0, 1

we further have

\| ∂^j (A_h J_h - J_h A) ξ \|_{L^∞ L^2} \lesssim h^k \| ξ \|_{H^{k+1}}, \quad (5.2a)

\| ∂^j A_h F_h - J_h ∂^j Λ F_h \|_{L^∞ L^2} \lesssim h^k \| ∂^j (λ, f) \|_{H^{k+1}}. \quad (5.2b)

Proof. Using the identity

L^V_h A^{-1} \psi = A^{-1} (L^V_h - L^H_h) \psi,

we compute

(T_h J_h - J_h T) ξ = \left( A^{-1} (L^V_h - L^H_h) \psi \right).

Using Lemma 2.7, we obtain

\| A^{-1} (L^V_h - L^H_h) \psi \|_{L^∞ L^2} \lesssim \| (L^V_h - L^H_h) \psi \|_{L^2},

\| A^{-1} (L^V_h - L^H_h) \psi \|_{L^∞ L^2} \lesssim \| (L^V_h - L^H_h) \psi \|_{V_h}.\n
and the bound (5.1a) and (5.1b) follow from (2.10) and (2.11). For the bound on (5.1c) we exploit the bounds on (2.8), (2.9), (2.11), (2.13), and (2.14) to obtain
\[
\| (\Lambda_h J_h - J_h \lambda) \psi \|_{L^\infty \times L^4} \lesssim \| \lambda \partial_t \psi - \mathcal{L}_h^V \psi \|_{H^1} + \| \pi_h (\lambda \mathcal{L}_h^V \psi - \mathcal{L}_h^{-1} \lambda \psi) \|_{H^1} + \| \pi_h (\mathcal{L}_h^V \psi - \mathcal{L}_h^{H^k}) \lambda \psi \|_{H^1} + \| (\mathcal{L}_h^{H^k} - \mathcal{L}_h^V) \lambda \psi \|_{H^1} + \| (\mathcal{L}_h^{H^k} - \mathcal{L}_h^V) \lambda \psi \|_{H^1
}\]
\[
(5.3)
\]
and in the same way for \( j = 0, 1 \)
\[
\| \partial_t^j (\Lambda_h J_h - J_h \lambda) \psi \|_{H^1 \times L^2} \lesssim h^k (\| \psi \|_{H^1} + \| (\partial_t^j \lambda) \psi \|_{H^1}).
\]
Similarly, (5.2b) follows from
\[
\| \partial_t^j \Lambda F - J_h \partial_t^j AF \|_{L^\infty \times L^4} \lesssim \| \partial_t^j \pi_h (\lambda \partial_t \psi - \mathcal{L}_h^{-1} \lambda \psi) \|_{L^4} + \| \partial_t^j (\pi_h \mathcal{L}_h^{-1} - \mathcal{L}_h^{H^k}) \lambda \psi \|_{L^4} + \| (\mathcal{L}_h^{H^k} - \mathcal{L}_h^V) \partial_t^j \lambda \psi \|_{L^4}
\]
by (2.8), (2.9), (2.11), (2.13), and (2.14).

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 \( u \in C^3([0, T], H^{k+1}(\Omega)) \cap C^4([0, T], H^{k+1}(\Omega)) \) and \( \lambda, f \in C^1([0, T], H^{k+1}(\Omega)) \).

(a) The defect \( \delta_{h,T}(t) \) introduced in (3.4) satisfies for \( t \in [0, T] \)
\[
\| \delta_{h,T}(t) \|_{W^{1, \infty} \times L^4} + \| \partial_t \delta_{h,T}(t) \|_{L^\infty \times L^4} \lesssim h^k.
\]
(b) If in addition \( u \in C^2([0, T], W^{1, \infty}(\Omega)) \cap C^4([0, T], H^1(\Omega)) \), then the defect \( \delta_{h,T}^u \) introduced in (4.2) satisfies for \( n \geq 2 \)
\[
\| \delta_{h,T}^u \|_{W^{1, \infty} \times L^4} + \| \partial_t \delta_{h,T}^u \|_{L^\infty \times L^4} \lesssim \tau + h^k.
\]

**Proof.** (a) We decompose the defect \( \delta_{h,T} = \delta_{h,T}^1 + \delta_{h,T}^2 \) introduced in (3.4) with
\[
\delta_{h,T}^1 = (\Lambda_h J_h - J_h \Lambda) \partial_t y, \quad \delta_{h,T}^2 = J_h T A F - T_h \Lambda_h F_h,
\]
and first consider \( \delta_{h,T}^1 \). We write
\[
\| \delta_{h,T}^1 \|_{W^{1, \infty} \times L^4} \leq \| T_h (\Lambda_h J_h - J_h \Lambda) \partial_t y \|_{W^{1, \infty} \times L^4} + \| (T_h J_h - J_h T) \Lambda_x y \|_{W^{1, \infty} \times L^4} + \| \Lambda \partial_t^2 u \|_{H^{k+1}}
\]
\[
\lesssim \| (\Lambda_h J_h - J_h \Lambda) \partial_t y \|_{L^\infty \times L^4} + \| \Lambda \partial_t^2 u \|_{H^{k+1}},
\]
where we used Lemma 2.7 together with (5.1a) and (5.1c). Differentiating and taking norms, gives with Lemma 2.3
\[
\| \partial_t \delta_{h,T}^1 \|_{L^\infty \times L^4} \leq \| T_h (\Lambda_h J_h - J_h \Lambda) \partial_t^3 y \|_{L^\infty \times L^4} + \| T_h (\partial_t \Lambda_h J_h - J_h \partial_t \Lambda) \partial_t y \|_{L^\infty \times L^4} + \| (T_h J_h - J_h T) \partial_t (\Lambda_x y) \|_{L^\infty \times L^4}
\]
\[
\lesssim \| (\Lambda_h J_h - J_h \Lambda) \partial_t^3 y \|_{H^{1 \times L^2}} + \| (\partial_t \Lambda_h J_h - J_h \partial_t \Lambda) \partial_t y \|_{H^{1 \times L^2}} + \| (T_h J_h - J_h T) \partial_t (\Lambda_x y) \|_{L^\infty \times L^4}.
\]
such that (5.1b) and (5.2a) together imply
\[ \| \partial_t \delta^1 h \|^2 \leq h^k (\| \partial_t \phi \|^2 + \| \partial_t^2 u \|^2 + \| \partial_t (\lambda \partial_t u) \|^2). \]

With similar arguments, we obtain for \( j = 0, 1 \) with (5.1a), Lemma 2.7, and (5.2b)
\[ \| \partial_t \delta^j h \|^2 \leq \| (J_0 - T_h J_h) \partial_t (A(t) F(t)) \|_{W_1,-X,L^1} + \| T_h (J_h \partial_t (A(t) F(t)) - \partial_t (A(t) F(t))) \|_{W_1,-X,L^1} \]
\[ \lesssim h^k \| \partial_t (A(t) F(t)) \|_{H^{k+1}}. \]

and thus the claim.

(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
\[ \| \partial_t \delta^j h \| \leq \| \delta_t y(t^n) - \partial_t y(t^n) \|_{W_1,-X,L^1} \]
\[ \| \partial_t \delta^j h \| \leq \| \delta_t y(t^n) - \partial_t y(t^n) \|_{H^{1},X,L^1} \]
\[ \| \partial_t \delta^j h \| \leq \| \delta_t y(t^n) - \partial_t y(t^n) \|_{H^{1},X,L^1} + \| \partial_t (\delta_t y(t^n) - \partial_t y(t^n)) \|_{H^{1},X,L^1}. \]

We note the identity
\[ \partial_t y(t^n) - \partial_t y(t^n) = \frac{1}{t^n} \int_{0}^{1} (-s) \partial_t^2 y(t^n) + ts \) ds, \]
and conclude the assertion. \( \square \)

In the next step, we consider the time derivatives of the error which we estimate in the energy norm, and provide some more approximation properties. We first characterize the inverse of \( \lambda_h \), which plays a crucial role in the subsequent error analysis. We define the inner product \( (\varphi_h | \psi_h)_{\lambda_h} := (J_0 \lambda \varphi_h | \psi_h)_{L^2(\Omega_h)} \), and the corresponding \( L^2 \)-projection \( Q_h \) for \( \psi \in L^2(\Omega_h) \) via
\[ (Q_h \psi | \psi_h)_{\lambda_h} = (\psi | \psi_h)_{\lambda_h}, \quad \psi_h \in V_h. \]

This leads us to the following result, which is proved in Appendix A.

**Lemma 5.3** (a) For \( \varphi \in H^{k+1}(\Omega) \) and \( \psi \in H^1(\Omega_h) \), we have the bounds
\[ \| Q_h \varphi \|_{L^2(\Omega_h)} \lesssim \| \varphi \|_{H^{k+1}(\Omega)}, \quad \| Q_h \psi \|_{H^1(\Omega_h)} \lesssim \| \psi \|_{H^1(\Omega_h)}. \]

(b) The inverse of \( \lambda_h \) is given by \( \lambda_h^{-1} \varphi_h = Q_h (J_0 \lambda^{-1} \varphi_h) \) for \( \varphi_h \in V_h \).

We further make use in the following of the operators
\[ \bar{\lambda}(t^n) := \begin{pmatrix} \lambda(t^n) & 0 \\ 0 & \text{Id} \end{pmatrix} \]
and analogously defined, \( \bar{\lambda}_h (t^n), \bar{\lambda}_h^{-1}(t^n) \), and \( \bar{\lambda}(t^n)^{-1} \).

**Lemma 5.4** Let \( \xi = (\varphi, \psi) \in H^{k+1}(\Omega) \times H^k(\Omega) \) and \( \lambda, f \in C^1([0,T], H^k(\Omega)). \) Then, the discrete operators introduced in Section 2 satisfy for \( j = 0, 1 \) the bounds
\[ \| \bar{\lambda}_h J_h - J_h \bar{\lambda} \| \lesssim h^k \| \varphi \|_{H^{k+1}}, \]
\[ \| \bar{\lambda}_h^{-1} J_h - J_h \bar{\lambda}^{-1} \| \lesssim h^k \| \varphi \|_{H^{k+1}}, \]
\[ \| \partial_t \bar{\lambda}_h F_h - J_0 \partial_t A F \| \lesssim h^k \| \partial_t (\lambda f) \|_{H^{k+1}}. \]
If \( \xi = (\varphi, \psi) \in (H^{k+3}(\Omega) \cap \mathcal{D}(\Delta^2)) \times (H^{k+2}(\Omega) \cap \mathcal{D}(\Delta)) \) and \( f \in C([0,T], H^{k+1}(\Omega) \cap \mathcal{D}(\Delta)) \), then
\[
\begin{align*}
\|A_h F_h - J_h AF\|_{X_h} &\lesssim h^k \|f\|_{H^{k+1}}, \\
\| (A_h J_h - J_h A) \xi \|_{X_h} &\lesssim h^k \|\Delta \varphi\|_{H^{k+1}}, \\
\| (A_h^2 J_h - J_h A^2) \xi \|_{X_h} &\lesssim h^k \|\Delta \xi\|_{H^{k+3} \times H^{k+1}}.
\end{align*}
\]

**Proof.** The same computation as in (5.3) yields (5.7a) and with Lemma 5.3 also (5.7b), and the representation in (5.4) implies (5.7c). We obtain
\[
\|A_h F_h - J_h AF\|_{X_h} = \|f - \mathcal{L}_h^{V^*} f\|_{V_h},
\]
and thus with (2.11) and (2.8) the claim. We use the identity \( \Delta_h \mathcal{L}_h^{V^*} = \mathcal{L}_h^{H^*} \Delta_h \) to derive
\[
(A_h J_h - J_h A) \xi = \left( \mathcal{L}_h^{H^*} - \mathcal{L}_h^{V^*} \right) \Delta \varphi,
\]
\[
(A_h^2 J_h - J_h A^2) \xi = \left( \mathcal{L}_h^{H^*} - \mathcal{L}_h^{V^*} \right) \Delta \psi.
\]

We employ (2.10) and (2.11) to obtain (5.8b). Together with the inverse estimate (2.17), this further implies (5.8c).

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

**Lemma 5.5** Let the solution \( u \in C^4([0,T], H^k(\Omega)) \cap C^2([0,T], \mathcal{D}(\Delta)) \) and \( \lambda, f \in C^2([0,T], H^k(\Omega)) \).

(a) The defect defined in (3.6) satisfies for \( j = 1, 2 \) and \( t \in [0,T] \)
\[
\| \partial_j^t \delta_h(t) \|_{X_h} \lesssim h^k.
\]

(b) If in addition \( u \in C^3([0,T], H^1(\Omega)) \), then the defect defined in (4.4) satisfies for \( j = 1, 2 \) and \( n \geq j \)
\[
\| \partial_j^t \delta_h^o(t) \|_{X_h} \lesssim \tau + h^k.
\]

**Proof.** (a) We estimate the defect with the help of Lemma 5.4. We employ Lemma 2.6 and the bounds (5.2a), (5.7c) and (5.8b) to obtain
\[
\| \partial_j^t \delta_h(t) \|_{X_h} \lesssim C h^k \left( \sum_{l=2}^4 \| \partial_l^j u \|_{H^{l+1}} + \| \partial_l^j \Delta u \|_{H^{l+1}} + \| \partial_l^j (\lambda f) \|_{H^{l+1}} \right).
\]

(b) Thanks to Lemma 4.3, it remains to show for \( j = 1, 2 \)
\[
\| \partial_j^t \delta_h(t^n) J_h (\partial_\tau^j y(t^n) - \partial_\tau^j y(t^n)) \|_{X_h} \lesssim \tau \| \partial_j^t \delta_h(t^n) \|_{H^1 \times H^1}.
\]

Using the continuity of \( J_h \) in \( H^1 \times H^1 \) and the identity (5.5) 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 (2.21) where \( y_h^0 = y_h(0) = J_h y_0 \). The aim of this section is to prove the following bounds.
LEMMA 5.6 Under the assumptions of Theorem 2.8 it holds for $\ell = 1, 2$

$$\|\partial^\ell_x e_h(0)\|_{X_h} \leq Ch^k.$$  

LEMMA 5.7 Under the assumptions of Theorem 2.9 it holds for $\ell = 1, 2$

$$\|\partial^\ell_x e_h^\ell\|_{X_h} \leq C(\tau + h^k).$$

In order to conclude the desired assertion, we proceed in a series of lemmas and introduce the notation $A^n = A(t^n), F(t^n) = F(t^n), A_h^n = A_h(t^n), F_h(t^n) = F_h(t^n)$ and also

$$\mathcal{B}^n = (A^n - \tau A)^{-1}, \quad \mathcal{B}_h^n = (A_h^n - \tau A_h)^{-1}. \quad (5.9)$$

We provide a detailed proof of Lemma 5.7 first and explain afterwards how to conclude the assertion of Lemma 5.6. In order to keep the notation simple, we assume without loss of generality in the following that the spatial order satisfies

$$k^* \leq \ell_{\text{max}}, \quad (5.10)$$

with $\ell_{\text{max}}$ defined in Assumption 2.1. This induces the restriction $\ell_{\text{max}} \geq 2$. In a first step, we find a suitable representation for the discrete and exact solutions.

LEMMA 5.8 The numerical solution can be expanded via

$$\begin{align*}
\partial_t y_h^n &= \mathcal{B}_h^n A_h y_h^{n-1} + G_h^n, \\
\partial_t^2 y_h^n &= \mathcal{B}_h^n A_h \mathcal{B}_h^{n-1} A_h y_h^{n-2} + (\partial_t \mathcal{B}_h^n) A_h y_h^{n-2} + \mathcal{B}_h^n A_h G_h^{n-1} + \partial_t G_h^n.
\end{align*}$$

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

$$G_h^n \to G^n = \mathcal{B}^n A^n (F^n + \delta^n), \quad \delta^n = \partial_t y(t^n) - \partial_t y(t^n). \quad (5.11)$$

Proof. Starting from (2.24), we multiply by $A_h^n$ and reorder the terms to obtain

$$(A_h^n - \tau A_h) y_h^n = A_h^n y_h^{n-1} + \tau A_h^n F_h^n.$$  

Using the resolvent, this further gives

$$y_h^n = y_h^{n-1} + \tau \mathcal{B}_h^n A_h y_h^{n-1} + \tau \mathcal{B}_h^n A_h F_h^n,$$

which implies the representation for $\partial_t y_h^n$. The remaining identity is deduced from the product rule in Lemma 4.1. The results for the exact solution can be derived from the representation

$$A^n \partial_t y(t^n) = A y(t^n) + A^n F^n + A^n \delta^n$$

and the same computations as above. \qed

In order to bound the expressions in Lemma 5.7, we subtract the representations for the exact and the numerical solution. Let us for example consider the first term in $\partial_t^2 e_h^2 = J_h \partial_t^2 y(t^2) - \partial_t^2 y_h^2$ given by

$$(\partial_t^2 e_h^2)_1 = \left( J_h \mathcal{B}^2 A \mathcal{B}_h^1 A - \mathcal{B}_h^1 A_h \mathcal{B}_h^1 A_h J_h \right) y^0,$$
where we used the initial value (2.21) with $y_h^0 = J_h y^0$. 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

$$\Lambda \mathcal{R}^n A^n = \Lambda^n \mathcal{R}^n A, \quad A \Lambda^n = \tilde{\Lambda}^n A,$$

(5.12)

with $\tilde{\Lambda}$ defined in (5.6). The corresponding equalities are also valid for the discrete objects and the inverse $A(A^n)^{-1} = (\tilde{\Lambda}^n)^{-1} A$. Hence, we can write

$$J_h \mathcal{R}^2 \Lambda \mathcal{R}^1 A y^0 = J_h \mathcal{R}^2 \tilde{\Lambda}^1 (\tilde{\Lambda}^1)^{-1} A^2 y^0$$

and similarly for the discrete counterpart. A reformulation of Lemma 5.8 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 (5.8). The remaining differences in front of them are treated by the following abstract estimate. As a shorthand notation, we set

$$\prod_{j=1}^m B^j := B^m \ldots B^1, \quad m \geq 1, \quad \text{and} \quad \prod_{j=1}^m B^j := \text{Id}, \quad m < 1.$$

Lemma 5.9 Let $\mathcal{Y} \subseteq X$ be a Hilbert space, $m \in \mathbb{N}$, and consider operators $B^j_h \in \mathcal{L}(X_h)$ and $B^j \in \mathcal{L}(X)$, $j = 1, \ldots, m$, with the following properties:

$$\| (J_h B^j - B^j_h) x \|_{X_h} \lesssim (\tau + h^k) \| x \|_{\mathcal{Y}},$$

$$\| B^j x \|_{\mathcal{Y}} \lesssim \| x \|_{\mathcal{Y}}.$$

(5.13)

Then, the product is bounded by

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

Proof. Using the telescopic sum

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

we immediately conclude the assertion. \hfill \square

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

Proof of Lemma 5.7. 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

$$B^j \in \{ A^n, \partial_x A^n, (\tilde{\Lambda}^n)^{-1}, \mathcal{R}^n \}$$

(5.14)

and the discrete counterparts using Lemma 5.9 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 $\mathcal{D}(A^k) \hookrightarrow H^{k+1}(\Omega) \times H^k(\Omega)$, see, e.g., (Grisvard, 1985, Rem. 2.5.1.2).
(i) We first consider the operators involving $A$. Under Assumption 2.1 for $k = k + 1$ and with Lemmas 2.3, 2.6, and 5.3 and the bounds (5.7a) and (5.7b), the properties (5.13) are satisfied. For the resolvent $\mathcal{R}^n$, we directly employ Lemma A.2 to compute with (5.2a) and (5.8b)

\[
\| (J_h \mathcal{R} - \mathcal{R}^n J_h) y \|_{X_h} \leq \tau \| J_h (A - A_h) \mathcal{R} y \|_{X_h} + \| J_h (A - A_h) \mathcal{R} y \|_{X_h} \\
\leq C h^k \tau \| A^{k+1} \mathcal{R} y \|_{X_h} + C h^k \| A^k \mathcal{R} y \|_{X_h},
\]

and hence Lemma 5.9 is applicable.

(ii) We employ Lemma A.1 and denote any product of operators from part (i) by $\Pi$ and the discrete counterpart $\Pi_h$. 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

\[
x \in \{ A y, A^2 y, F^1, \partial \tau F^2, A F^1 \}, \quad x_h \in \{ A_h J_h y, A_h^2 J_h y, F^1_h, \partial \tau F^2_h, A_h F^1_h \},
\]

(5.15) and $\delta \in \{ \delta^1, \partial \tau \delta^2, A \delta^1 \}$ with $\delta^i$ given in (5.11). Then, the first part is covered by part (i), provided that $\| A^k x \|_{X} \leq C$ holds for all $x$ in (5.15), which follows from the assumptions of the theorem. The second part is bounded due to (5.7c), (5.8a), (5.8b), and (5.8c). Lastly, the estimate

\[
\| J_h \Pi \delta \|_{X_h} \leq A \delta \|_{X}
\]

together with (5.5) yields

\[
\| J_h \Pi \delta \|_{X_h} \leq \tau ( \| A^2 \partial \tau^3 y \|_{L^2(X)} + \| A \partial \tau^3 y \|_{L^2(X)} ),
\]

and hence the desired bound. \hfill \square

**Remark 5.10** The restriction $\ell_{\text{max}} \geq k^*$ in (5.10) is necessary in order to guarantee (2.5) for $\ell = k^*$ which is needed in the estimates above to transport the regularity past $A$.

We now prove Lemma 5.6 and observe that formally setting $\tau = 0$ in (2.24), we obtain the spatially discretized equation (2.20). This observation drives the following proof.

**Proof of Lemma 5.6.** Several steps in the proof simplify, and we mainly explain why we can neglect the assumption on $\ell_{\text{max}}$. First note, that compared to the proof above we replace due to $\tau = 0$

\[
\mathcal{R}^n \rightarrow (A^0)^{-1}, \quad \mathcal{R}_h^n \rightarrow (A_h^0)^{-1},
\]

and the most delicate term is given by

\[
(\partial^2 \varepsilon h(0))_1 = (J_h (A^0)^{-1} A (A^0)^{-1} - (A_h^0)^{-1} A_h (A_h^0)^{-1} A_h) y h^0
\]

\[
= (J_h (A^0)^{-1} (A^0)^{-1} A^2 - (A_h^0)^{-1} (A_h^0)^{-1} A_h^2 J_h) y^0,
\]

Further, the terms in (5.14) reduce to $B^j \in \{ A^0, \partial h A^0, (A^0)^{-1} \}$. In contrast to above, we may employ Lemma 5.9 with $\mathcal{W} = H^{k+1}(\Omega) \times H^k(\Omega)$, and thus obtain no restriction on $\ell_{\text{max}}$. Then, we proceed along the lines of Lemma 5.7. \hfill \square
6. Numerical experiments

To illustrate our theoretical findings, we present some numerical experiments. Let \( \Omega = B_1(0) \subset \mathbb{R}^2 \) be the two-dimensional unit sphere and consider equation (1.1) with data given by

\[
\begin{align*}
    u_0(x) & = \frac{1}{20} \sin(\pi r^2) - 7 , \\
    v_0(x) & = \frac{1}{20} \sin(\pi r^2) - 7 , \\
    \lambda(t, x) & = 2 + (1 - r^2) x_1 e^t + (1 - r^2) x_2 \cos(t).
\end{align*}
\]

where \( r^2 = |x|^2 \). The right-hand side \( f \) is chosen such that the exact solution is given by

\[
    u(t, x) = \frac{1}{20} e^t \sin(\pi r^2) - 7.
\]

A simple calculation shows, that the regularity assumptions of Theorems 2.8 and 2.9 are satisfied with \( k = 3 \) and \( m_{\mu, \max} = 4 \). The scaling by a factor 20 is used to approximately normalize the \( W_1^1, \infty \)-norm of the solution \( u \).

6.1 Discretization

To discretize in space, we multiply (1.1) with \( \lambda(t, x) \) and obtain the equivalent formulation

\[
    \lambda(t, x) \partial_{tt} u(t, x) = \Delta u(t, x) + \lambda(t, x) f(t, x).
\]

Using the mass and stiffness matrix defined by

\[
    M_h(t_i)_{i,j} = \left( \phi_i | \phi_j \right)_{L^2(\Omega_h)} , \quad L_h_{i,j} = \left( \nabla \phi_i | \nabla \phi_j \right)_{L^2(\Omega_h)},
\]

where we denote by \( (\phi_i)_{i} \) the nodal basis of \( V_h \), the discrete solution in (2.20) satisfies

\[
    M_h(t_n) \partial_{tt} u_h(t) = -L_h u_h(t) + M_h(t_n) I_h f(t),
\]

by abusing the notation for the coefficient vectors and their corresponding function in \( V_h \). The fully discrete method in (2.24) is then given for \( n \geq 2 \) by

\[
    M_h(t_n) u_h(t) - 2 u_h(t) + u_h(t-2) = -\tau_2 L_h u_h(t) + \tau_2 M_h(t_n) I_h f(t),
\]

and for \( n = 1 \), we have to solve

\[
    M_h(t_1) u_h(t) - u_0 - \tau_1 v_0 = -\tau_2 L_h u_h(t) + \tau_2 M_h(t_1) I_h f(t).
\]

Further, we have \( v_n = \partial_{\tau} u_n \) for \( n \geq 1 \). We implemented the numerical experiments in \( C++ \) using the finite element library \( \text{deal.II} \) (version 9.4) Arndt et al. (2022); Bangerth et al. (2007). A precise description of the implementation can for example be found in (Leibold, 2021, Ch. 6.5.1). The codes written by Malik Scheifinger to reproduce the experiments are available at \( \text{https://doi.org/10.5445/IR/1000157919} \).
6. Numerical experiments

To illustrate our theoretical findings, we present some numerical experiments. Let \( \Omega = B_1(0) \subset \mathbb{R}^2 \) be the two-dimensional unit sphere and consider equation (1.1) with data given by

\[
\begin{align*}
    u^0(x) &= \frac{1}{20} \sin(\pi r^2)^7, \quad v^0(x) = \frac{1}{20} \sin(\pi r^2)^7, \\
    \lambda(t,x) &= 2 + (1 - r^2)^2 (x_1 \psi + (1 - r^2)x_2 \cos(t)),
\end{align*}
\]

where \( r^2 = |x|^2 \). The right-hand side \( f \) is chosen such that the exact solution is given by

\[
u(t,x) = \frac{1}{20} \psi \sin(\pi r^2)^7.
\]

A simple calculation shows, that the regularity assumptions of Theorems 2.8 and 2.9 are satisfied with \( k = 3 \) and \( \ell_{\max} = 4 \). The scaling by a factor 20 is used to approximately normalize the \( W^{1,\infty} \)-norm of the solution \( u \).

6.1 Discretization

To discretize in space, we multiply (1.1) with \( \lambda(t,x) \) and obtain the equivalent formulation

\[
\lambda(t,x) \partial_t u(t,x) = \Delta u(t,x) + \lambda(t,x) f(t,x).
\]

Using the mass and stiffness matrix defined by

\[
(M_h(t))_{i,j} := ((I_h \lambda(t)) \varphi_i | \varphi_j)_{L^2(\Omega_h)}, \quad (L_h)_{i,j} := (\nabla \varphi_i | \nabla \varphi_j)_{L^2(\Omega_h)},
\]

where we denote by \( (\varphi_i) \) the nodal basis of \( V_h \), the discrete solution in (2.20) satisfies

\[
M_h(t) \partial_t u_h(t) = -L_h u_h(t) + M_h(t) I_h f(t),
\]

by abusing the notation for the coefficient vectors and their corresponding function in \( V_h \). The fully discrete method in (2.24) is then given for \( n \geq 2 \) by

\[
M_h(t^n) (u^n_h - 2u^{n-1}_h + u^{n-2}_h) = -\tau^2 L_h u^n_h + \tau^2 M_h(t^n) I_h f(t^n),
\]

and for \( n = 1 \), we have to solve

\[
M_h(t^1) (u^1_h - u^0_h - \tau v^0_h) = -\tau^2 L_h u^1_h + \tau^2 M_h(t^1) I_h f(t^1).
\]

Further, we have \( v^n_h = \partial_t u^n_h \) for \( n \geq 1 \). We implemented the numerical experiments in C++ using the finite element library deal.II (version 9.4) Arndt et al. (2022); Bangerth et al. (2007). A precise description of the implementation can for example be found in Leibold, 2021, Ch. 6.5.1. The codes written by Malik Scheifinger to reproduce the experiments are available at https://doi.org/10.5445/IR/1000157919.
6.2 Numerical results

For the problem described above, we performed experiments for the time and space discretization. To this end, we used finite elements of order \( k = 1, 2, 3 \). Since the computation of the lift of a finite element function is very laborious, we do not compute the full error in the form \( \mathcal{L}_h u - u_h \). Instead, in our numerical examples we consider the error

\[
E(t) := \| u_h(t) - I_h u(t) \|_{W^{1,-}(\Omega_h)} + \| v_h(t) - I_h \partial_t u(t) \|_{L^-}(\Omega_h),
\]

which is of the same order by the standard interpolation estimates. In the left part of Figure 3, the convergence of the error with respect to the spatial mesh width \( h \) is shown when using the implicit Euler method with \( \tau_{ref} = 8 \cdot 10^{-5} \). We observe that for finite elements of order \( k \) the error converges with order \( k \) in space as predicted by Theorem 2.8 until the error for \( k = 3 \) is dominated by the error of the temporal approximation. In the right part of Figure 3, we consider the convergence of the error with respect to the time step size \( \tau \). In space, we discretized with finite elements of order 3 and \( h_{ref} = 1.52 \cdot 10^{-2} \) such that the spatial error is negligible. Aligning to Theorem 2.9, we observe convergence of order 1 in time.

Acknowledgment

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 258734477 – SFB 1173. 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. Further, we like to thank Malik Schefinger for providing us the codes for the numerical experiments.

REFERENCES


A. Appendix

In this appendix, we provide the proof of Lemma 2.7, properties (2.12) and (2.13), and the postponed calculations from Section 5. The following proof is adapted from the conforming case presented in (Bramble et al., 1977, Lem. 4.1), and Dörich (2022).

Proof of Lemma 2.7. We first recall the solution operators $\Delta^{-1}$ and $\Delta_{h}^{-1}$ defined in (2.6) and (2.19), respectively. We further define the modified solution operator $\bar{S}_h = \mathcal{L}_h^{1/2} \Delta^{-1} \mathcal{L}_h$, and use it to expand $\Delta_h^{-1} = \bar{S}_h + (\Delta_h^{-1} - \bar{S}_h)$. For the first term, we use Sobolev’s embedding, the stability of the Ritz map in $W^{1,p}$, $2 \leq p \leq \infty$, (which is interpolated from (2.11) and (2.15)), as well as Theorem 2.4 to obtain

$$
\| \bar{S}_h \phi_h \|_{L^2(\Omega_h)} \lesssim \| \bar{S}_h \phi_h \|_{W^{1,4} (\Omega_h)} \lesssim \| \Delta^{-1} \mathcal{L}_h \phi_h \|_{W^{1,4} (\Omega)} \lesssim \| \Delta^{-1} \mathcal{L}_h \phi_h \|_{H^2 (\Omega)} \lesssim \| \phi_h \|_{L^2 (\Omega_h)},
$$

and similarly using the same steps

$$
\| \bar{S}_h \phi_h \|_{W^{1,-} (\Omega_h)} \lesssim \| \Delta^{-1} \mathcal{L}_h \phi_h \|_{W^{1,-} (\Omega)} \lesssim \| \Delta^{-1} \mathcal{L}_h \phi_h \|_{W^{2,4} (\Omega)} \lesssim \| \mathcal{L}_h \phi_h \|_{L^4 (\Omega)} \lesssim \| \phi_h \|_{L^4 (\Omega)}.
$$

It remains to bound the difference, stemming from the nonconformity, by the inverse estimate (2.17) via

$$
\| \bar{S}_h \phi_h - \Delta_h^{-1} \phi_h \|_{W^{1,-} (\Omega_h)} \lesssim C h^{-N/2} \| \bar{S}_h \phi_h - \Delta_h^{-1} \phi_h \|_{V_h}
$$

$$
= C h^{-N/2} \sup_{\| \psi_h \|_{V_h} = 1} \left( \bar{S}_h \phi_h - \Delta_h^{-1} \phi_h , \psi_h \right)_{V_h}
$$

$$
= C h^{-N/2} \sup_{\| \psi_h \|_{V_h} = 1} \left( (\phi_h , \psi_h)_{L^2 (\Omega_h)} - (\mathcal{L}_h \phi_h , \mathcal{L}_h \psi_h)_{L^2 (\Omega)} \right),
$$

where we used (2.6), (2.9), and (2.19) in the last step. Finally, we employ (Elliott & Ranner, 2021, Lem. 8.24) to obtain

$$
\left( \phi_h , \psi_h \right)_{L^2 (\Omega_h)} - (\mathcal{L}_h \phi_h , \mathcal{L}_h \psi_h)_{L^2 (\Omega)} \lesssim h^{3/2} \| \phi_h \|_{L^2} \| \psi_h \|_{V_h},
$$

which yields the assertion. \qed

Next, we show the stability and convergence properties for $\mathcal{L}_h^{H^s}$ and $\pi_h$.

Proof of (2.12) and (2.13). Let $\phi_h \in V_h$, and estimate with (Elliott & Ranner, 2021, Lem. 8.24)

$$
\| \mathcal{L}_h^{H^s} \phi - \phi_h \|_{L^2} \lesssim h^{1/2} \| \phi_h \|_{L^2} + \| \phi - \mathcal{L}_h \phi_h \|_{L^2}
$$

and thus obtain with the inverse estimate (2.17)

$$
\| \mathcal{L}_h^{H^s} \phi \|_{V_h} \lesssim \| \phi_h \|_{V_h} + h^{-1} \| \mathcal{L}_h^{H^s} \phi - \phi_h \|_{L^2} \lesssim \| \phi_h \|_{V_h} + h^{-1} \| \phi - \mathcal{L}_h \phi_h \|_{L^2}.
$$

We now use the quasi-interpolation from (Bernardi, 1989, Thm 5.1), which is stable in $H^1$ and converges linearly in $L^2$, and hence gives (2.12).
Employing again (Elliott & Ranner, 2021, Lem. 8.24), for $\varphi \in H^2(\Omega)$ we compute
\[
| (\pi_h L^{-1}_h \varphi - L^{H^*} L^{-1}_h \varphi) , \psi_h |_{L^2(\Omega_h)} = | (L^{-1}_h \varphi , \psi_h)_{L^2(\Omega_h)} - (\varphi , L_h \psi_h)_{L^2(\Omega)} |
\leq C h^k \| \varphi \|_{L^2(\Omega_h)} \| L_h \psi_h \|_{L^2(\Omega)}
\leq C h^{k+1} \| \varphi \|_{H^2(\Omega)} \| L_h \psi_h \|_{L^2(\Omega)},
\]
where we used (Dörich, 2022, Appendix B) in the last step, and (2.13) is shown. □

Further, we verify the properties in Lemma 5.3 of the (weighted) $L^2$-projection $Q_h$ and the formula for the inverse of $\lambda_h$.

**Proof of Lemma 5.3.** (a) By the definition of $Q_h$, we have
\[
(L^{H^*} v - Q_h L^{-1}_h v | L^{H^*} v - Q_h L^{-1}_h v)_{\lambda_h} = (L^{H^*} v - Q_h L^{-1}_h v | L^{H^*} v - L^{-1}_h v)_{\lambda_h},
\]
and thus the equivalence of norms together with (2.8) and (2.10) gives the first claim. The stability in $H^1$ is shown as for (2.12).

(b) One easily verifies the identities for $\varphi_h, \psi_h \in V_h$
\[
(\lambda_h (\lambda_h^{-1} \varphi_h) , \psi_h)_{L^2(\Omega_h)} = (\varphi_h , \psi_h)_{L^2(\Omega_h)}, \quad (\lambda_h^{-1} (\lambda_h \varphi_h) , \psi_h)_{\lambda_h} = (\varphi_h , \psi_h)_{\lambda_h},
\]
which gives the last statement. □

Next, we give an extension of Lemma 5.8. We do not provide a proof here, since the expressions are derived by an iterative application of the identities (5.12).

**Lemma A.1** It holds the representation
\[
\partial_T y^{1}_h = R^1_h A_h M^0_h + R^1_h A_h (t^1) F^1_h, \\
\partial_T^2 y^{1}_h = R^2_h A_h^2 M^1_h - R^2_h \partial_T A_h^2 M^1_h + R^2_h A_h^2 \partial_T A_h^1 F^1_h + R^2_h A_h^2 \partial_T^2 F^1_h - R^2_h \partial_T A_h^2 A_h^1 F^1_h.
\]
The same expansion holds for $\partial_T y(t^\ell), \ell = 1, 2$, with $h$ formally set to zero and
\[
F^h_n \rightarrow F^n + \delta^n, \quad \delta^n = \partial_T y(t^n) - \partial_T y(t^n).
\]

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

**Lemma A.2** Let $R$ be the resolvent defined in (5.9) and $\ell_{\max}$ given in Assumption 2.1. Then, for $0 \leq \ell \leq \ell_{\max}$ there are constants $c_{\ell}$ such that
\[
\| A^\ell R y \|_X \leq C \| A^\ell y \|_X \\
\tau \| A^{\ell+1} R y \|_X \leq C \| A^\ell y \|_X,
\]
for all $y \in \mathcal{D}(A^{\ell_{\max}})$.

**Proof.** By the skew-adjointness of $A$ and (2.4), we derive
\[
\| R y \|_X \leq C \| y \|_X,
\]