Karlsruhe Institute of Technology

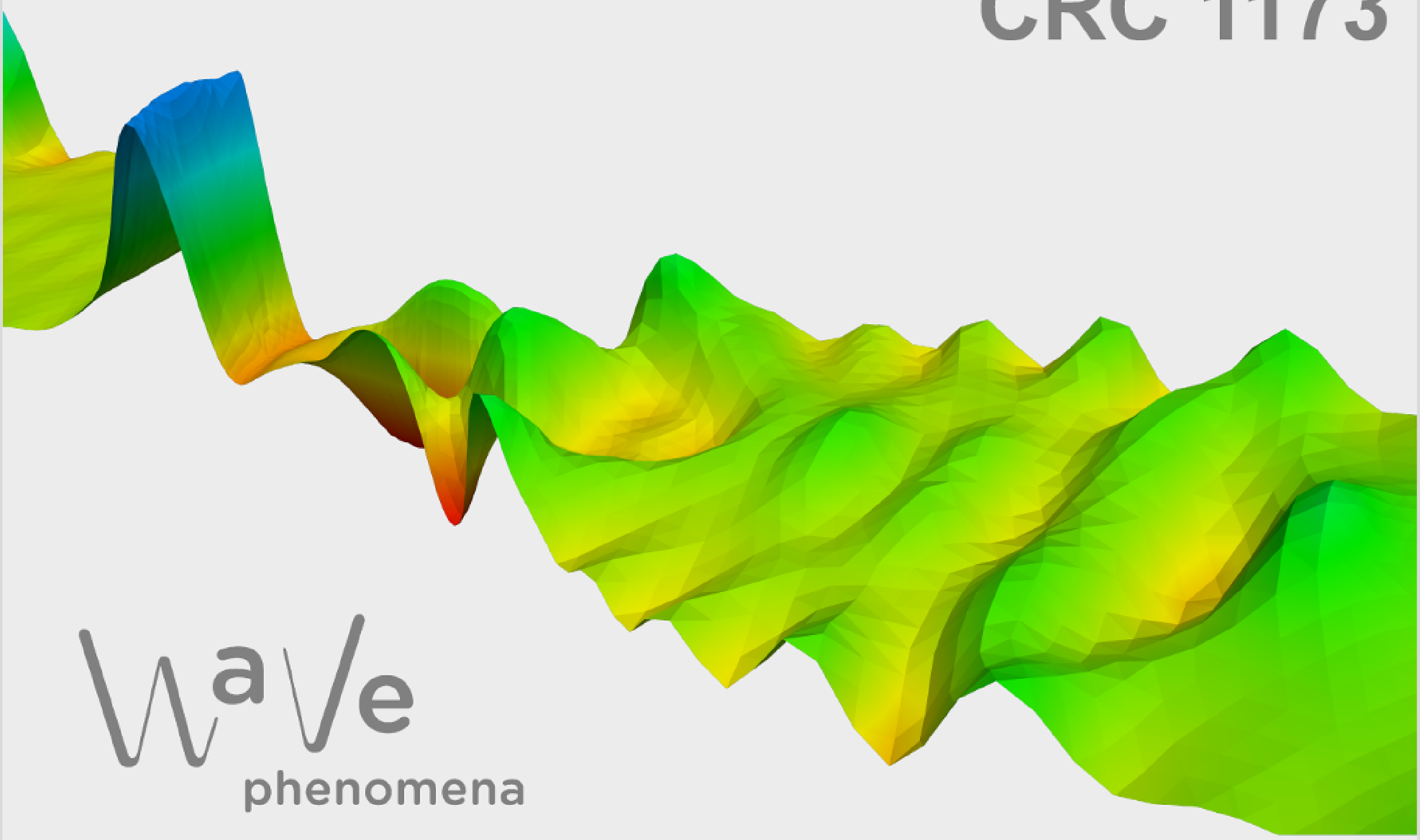
Strong norm error bounds for quasilinear wave equations under weak CFL-type conditions

Benjamin Dörich

CRC Preprint 2022/52 (revised), November 2023

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173



Participating universities

Universität Stuttgart

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

TU
WIEN

Funded by

DFG

ISSN 2365-662X

STRONG NORM ERROR BOUNDS FOR QUASILINEAR WAVE EQUATIONS UNDER WEAK CFL-TYPE CONDITIONS

BENJAMIN DÖRICH

ABSTRACT. In the present paper we consider a class of quasilinear wave equations on a smooth, bounded domain. We discretize it in space with isoparametric finite elements, and apply a semi-implicit Euler and midpoint rule as well as the exponential Euler and midpoint rule to obtain four fully discrete schemes. We derive rigorous error bounds of optimal order for the semi-discretization in space and the fully discrete methods in norms which are stronger than the classical $H^1 \times L^2$ energy norm under weak CFL-type conditions. To confirm our theoretical findings, we also present numerical experiments.

1. INTRODUCTION

In the present paper we consider the quasilinear wave equation

$$\lambda(u(t,x))\partial_{tt}u(t,x) = \Delta u(t,x) + g(t,x,u(t,x),\partial_t u(t,x)), \tag{1.1}$$

for $t \in [0,T]$, $x \in \Omega \subset \mathbb{R}^N$, $N = 1,2,3$. We assume the domain Ω to be bounded with a sufficiently regular boundary, and impose homogeneous Dirichlet boundary conditions. We discretize (1.1) in space using isoparametric finite elements, and employ for the time discretization a semi-implicit Euler and midpoint rule (which is in this case equivalent to the Crank–Nicolson scheme) as well as an exponential Euler and midpoint rule. We derive error bounds in norms stronger than the standard energy $H^1 \times L^2$-norm.

The first wellposedness results for a large class of quasilinear wave type equation was given by Kato in [25, 26]. This approach was refined in [11] for the problem (1.1) to account for the state-dependent norms necessary in the analysis. A typical example in nonlinear acoustics is the model $\lambda(u) = 1 - u^m$ for some $m \in \mathbb{N}$. Hence, in order to ensure $\lambda(u) > 0$, a key ingredient in the proof is to establish pointwise bounds on u (as well as $\partial_t u$), often via Sobolev's embedding $H^2 \hookrightarrow L^\infty$. To inherit this property in the spatial discretization, we need pointwise bounds on the numerical approximations in the error analysis. However, since the finite element space is not H^2-conforming, we cannot follow the above approach.

So far in the literature, bounds in $H^1 \times L^2$ are shown by inverse estimates which yield a factor $h^{-\beta}$ for some $\beta \geq 1$ with the spatial mesh width h. This induces unsatisfactory CFL-type conditions and excludes linear finite elements. In contrast to this, we adapt the idea from the wellposedness and perform the error analysis

2020 *Mathematics Subject Classification.* Primary: 65M12, 65M15. 65J15, Secondary: 65M60, 35L05.

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

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

not in the energy space $H^1 \times L^2$, but employ a discrete version of the H^2-norm. A discrete variant of Sobolev's embedding and a suitably defined solution space for the numerical approximation allow us to remove lower bounds on the polynomial degree of the finite element space and significantly improve the CFL-type condition compared to the literature. For the temporal step size τ and the spatial mesh width h, we show convergence in $N = 2$ under the restriction $\tau \lesssim h^\alpha$, for any $\alpha > 0$, and in $N = 3$ we have $\tau \lesssim h^{1/2+\alpha}$ for the first-order methods in time and $\tau \lesssim h^{1/4+\alpha}$ for the second-order method. In addition, we fully remove the CFL-type condition for $N = 1$.

The strategy of the semi discrete proof relies on a bootstrap argument. We set up a suitable solution space for the numerical approximation, and show that the initial value lies in this. Instead of the usual choice of interpolated initial values, we have to use a Ritz map for which we provide a computable alternative of the correct order. Since we are working with a finite dimensional subspace, this directly yields local wellposedness up to some time $t_h^* > 0$. On this possibly short time interval, we prove convergence in the stronger norm, and use this to extend t_h^* beyond T and to close the argument. For the fully discrete error bounds, this approach is generalized using an induction argument.

We give a brief overview of the literature on the numerical treatment of quasilinear wave equations. In the pioneering works [10, 24, 27, 37], existence of solutions to quasilinear and nonlinear evolution equations is established, and one can find approximation rates of the implicit and semi-implicit Euler method. Within an (extended) Kato framework, optimal order for these methods was achieved in [22] and rigorous error bounds for the time discretization by higher-order Runge-Kutta methods are derived in [23, 28].

Concerning the spatial discretization, the results in [21] yields optimal order of convergence for the equation (1.1), however only for polynomials of degree greater than two. For the strongly damped Westervelt equation, continuous and discontinuous Galerkin methods were analyzed in [1, 34]. Very recently, mixed finite elements for the Kuznetsov and Westervelt equations were studied in [33].

In [31], error bounds for two variant of the midpoint rule are derived of optimal order, but only for polynomials of degree greater than two and under a stronger CFL-type condition compared to our results. In the case of one-dimensional wave equation subject to periodic boundary conditions, full discretization error bounds are established in [19]. A sophisticated energy technique combined with the properties of the spectral discretization yields convergence without a CFL-type condition.

For a slightly different quasilinear wave equation, optimal error bounds in L^2 for continuous finite elements were considered in the literature. One-step methods of different order are analyzed in [3, 4, 17], and two-step methods are considered in [5]. For a class of linearly implicit single-step schemes as well as a linearly and a fully implicit two-step scheme, optimal error bounds are derived in [32]. However, all of these results require a CFL-type condition at least as strong as $\tau \lesssim h$, and do not allow for linear finite elements. We expect that our technique can be generalized to these problems, but this will be part of future research.

The paper is organized as follows: We describe in Section 2 the analytical framework and the space discretization by isoparametric Lagrange finite element, present the schemes and state our main results. We further show some numerical experiments to confirm our theoretical findings.

The proof of the spatial convergence rates is given in Section 3, where we first reduce the main result to error bounds in a stronger energy norm which is established afterwards. In Section 4, we extend this technique to the fully discrete case for the four presented methods. Certain stability estimates and the bounds on the defects are given in Section 5, and some postponed results are shown in Appendices A to D.

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 $a \leq Cb$, but it may depend on the polynomial degree k. For the sake of readability, we introduce the notation $t^n = n\tau$. 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 bounded domain $\Omega \subset \mathbb{R}^N$, $N = 1, 2, 3$, with boundary $\partial\Omega \in C^s$, $s \in \mathbb{N}$, we study the quasilinear wave equation (1.1) with homogeneous Dirichlet boundary conditions, and initial values

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

We note that the operator $-\Delta$ is positive and self-adjoint on $L^2(\Omega)$, and we define the spaces $H = L^2(\Omega)$ and $V = H^1_0(\Omega)$. Throughout the paper we impose the following conditions on the function λ and g. Additional requirements are stated in our main results.

Assumption 2.1. (λ_1) *The function $\lambda\colon \mathbb{R} \to \mathbb{R}$ satisfies $\lambda \in C^2(\mathbb{R}, \mathbb{R})$.*

(λ_2) *There is some radius $\widehat{r}_\infty > 0$ such that there is a constant $c_\lambda = c_\lambda(\widehat{r}_\infty) > 0$ such that*

$$c_\lambda \leq \lambda(x), \quad |x| \leq \widehat{r}_\infty.$$

(g_1) *The function $g\colon [0,T]\times\overline{\Omega}\times\mathbb{R}\times\mathbb{R} \to \mathbb{R}$ satisfies $g \in C^2([0,T]\times\overline{\Omega}\times\mathbb{R}\times\mathbb{R}, \mathbb{R})$.*

(g_2) *For $x \in \partial\Omega$ and $y = z = 0$ it holds $g(t, x, y, z) = 0$.*

The conditions (λ_1), (g_1) are structural assumptions which allow us to show crucial stability estimates. The lower bound in (λ_2) prevents the degeneracy of (1.1). The main effort in the discretization and error analysis is to ensure that this condition is inherited. We note that condition (g_2) implies in particular that for $u, v \in V$ one has $g(t, u, v) \in V$, and that all conditions are already required for the wellposedness. We recall an example for the quasilinear problem (1.1) given in [11].

Example 2.2. Let $K \in C^4(\mathbb{R}, \mathbb{R})$ with $1 + K'(0) > 0$ and consider the problem

$$\partial_{tt}(u + K(u)) = \Delta u,$$

for example with the Kerr model $K(z) = \alpha z^3$ for $\alpha \in \mathbb{R}$. If we rewrite it in the form (1.1), we obtain

$$\lambda(z) = 1 + K'(z), \qquad g(t, x, u, v) = -K''(u)v^2,$$

which satisfy Assumption 2.1. Denoting the fractional powers of the Laplacian by $\mathcal{H}_k \coloneqq \mathcal{D}\big((-\Delta)^{k/2}\big)$, under suitable smallness assumptions on the initial values, the existence of a solution

$$u \in C([0,T], \mathcal{H}_3) \cap C^1([0,T], \mathcal{H}_2) \cap C^2([0,T], \mathcal{H}_1)$$

is shown in [11, Thm. 4.1]. $\lozenge$

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

$$\Lambda(y(t))\partial_t y(t) = \mathrm{A}y(t) + G(t, y(t)), \qquad t \in [0, T], \qquad y = \begin{pmatrix} u \\ \partial_t u \end{pmatrix}, \tag{2.1}$$

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

$$y^0 = \begin{pmatrix} u^0 \\ v^0 \end{pmatrix}, \quad \Lambda(y) = \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & \lambda(u) \end{pmatrix}, \quad \mathrm{A} = \begin{pmatrix} 0 & \mathrm{Id} \\ \Delta & 0 \end{pmatrix}, \quad G(t, y) = \begin{pmatrix} 0 \\ g(t, u, \partial_t u) \end{pmatrix}.$$

Remark 2.3. The assumption on the regularity of the boundary is not essential in the error analysis, which also works on a convex, polygonal domain. Hence, one could apply a conforming finite element method. However, since the wellposedness of quasilinear equations requires a regular boundary, we will work in the nonconforming framework in the following.

Space discretization. We study the nonconforming space discretization of (2.1) based on isoparametric finite elements. For further details on this approach, we refer to [15, 16]. In particular, we introduce a shape-regular and quasi-uniform triangulation $\mathcal{T}_h$, consisting of isoparametric elements of degree $k \in \mathbb{N}$ and let $\partial\Omega \in C^{k+1}$. 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 $K \in \mathcal{T}_h$. In the following, we require that h is sufficiently small such that all cited results below hold true. We note that the smallness only depends on the geometry of the domain Ω and the polynomial degree k. The semi-discrete approximations are given by $u_h(t) \approx u(t)$ and $v_h(t) \approx \partial_t u(t)$. 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

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

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

$$H_h = \big(W_h, (\cdot \mid \cdot)_{L^2(\Omega_h)}\big), \quad V_h = \big(W_h, (\cdot \mid \cdot)_{H_0^1(\Omega_h)}\big)\,,$$

and we set $X_h = V_h \times H_h$.

Following the detailed construction in [16, 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

$$c_p \left\|\varphi_h\right\|_{L^p(\Omega_h)} \le \left\|\mathcal{L}_h\varphi_h\right\|_{L^p(\Omega)} \le C_p \left\|\varphi_h\right\|_{L^p(\Omega_h)}, \qquad \varphi_h \in L^p(\Omega_h), \tag{2.2a}$$

$$c_p \left\|\varphi_h\right\|_{W^{1,p}(\Omega_h)} \le \left\|\mathcal{L}_h\varphi_h\right\|_{W^{1,p}(\Omega)} \le C_p \left\|\varphi_h\right\|_{W^{1,p}(\Omega_h)}, \quad \varphi_h \in W^{1,p}(\Omega_h), \tag{2.2b}$$

cf. [16, Prop. 5.8]. By construction, the boundary nodes of Ω_h lie on $\partial\Omega$ and zero boundary conditions are preserved by $\mathcal{L}_h$, see [16, Sec. 8.5]. Further by [15, Sec. 4], the lift preserves values at the nodes, i.e., in particular

$$I_h\mathcal{L}_h\varphi_h = \varphi_h, \quad \varphi_h \in V_h\,, \tag{2.3}$$

where we denote the nodal interpolation operator by $I_h \colon C_0(\Omega) \to V_h$ and, enriching the space W_h by basis functions corresponding to the boundary nodes, its extension

$I_h^e\colon C(\Omega) \to C(\Omega_h)$. Further, 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

$$\big(\mathcal{L}_h^{H*}\varphi \mid \psi_h\big)_{H_h} = (\varphi \mid \mathcal{L}_h\psi_h)_{L^2(\Omega)}\,, \qquad \varphi \in H,\ \psi_h \in H_h, \tag{2.4a}$$

$$\big(\mathcal{L}_h^{V*}\varphi \mid \psi_h\big)_{V_h} = (\varphi \mid \mathcal{L}_h\psi_h)_{H_0^1(\Omega)}\,, \qquad \varphi \in V,\ \psi_h \in V_h. \tag{2.4b}$$

We note that in the conforming case, i.e. $\Omega = \Omega_h$, $\mathcal{L}_h^{H*}$ and $\mathcal{L}_h^{V*}$ coincide with the L^2-projection $\pi_h\colon L^2(\Omega_h) \to H_h$ and the Ritz projection $R_h\colon H_0^1(\Omega_h) \to V_h$, respectively, given by

$$(\pi_h\psi \mid \psi_h)_{H_h} = (\psi \mid \psi_h)_{L^2(\Omega_h)}\,, \qquad \psi \in L^2(\Omega_h),\ \psi_h \in H_h, \tag{2.5a}$$

$$(R_h\psi \mid \psi_h)_{V_h} = (\psi \mid \psi_h)_{H_0^1(\Omega_h)}\,, \qquad \psi \in H_0^1(\Omega_h),\ \psi_h \in V_h. \tag{2.5b}$$

For $u_h, v_h \in V_h$ we define the discrete operator $\lambda_h(u_h)\colon H_h \to H_h$ and the discrete right-hand side g_h by

$$\lambda_h(u_h)\varphi_h = \pi_h\big(I_h\lambda(\mathcal{L}_h u_h)\,\varphi_h\big), \qquad g_h(t,u_h,v_h) = I_h g(t,\mathcal{L}_h u_h,\mathcal{L}_h v_h), \tag{2.6}$$

respectively. Denoting by $I_h^{\Omega_h}\colon C(\Omega_h) \to V_h$ the nodal interpolation operator with the same nodes as I_h, we have by (2.3) the identity $I_h\lambda(\mathcal{L}_h u_h) = I_h^{\Omega_h}\lambda(u_h)$, and similarly for g_h. The first-order counterparts of (2.6) are given by

$$y_h = \begin{pmatrix} u_h \\ v_h \end{pmatrix}, \quad \Lambda_h(y_h) = \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & \lambda_h(u_h) \end{pmatrix}, \quad G_h(t,y_h) = \begin{pmatrix} 0 \\ g_h(t,u_h,v_h) \end{pmatrix}. \tag{2.7}$$

Moreover, we show below in Section 2.5 that under certain assumption on $u_h \in V_h$ there exists a modified L^2-projection $Q_h(u_h)\colon L^2(\Omega_h) \to V_h$ such that the inverse of $\lambda(u_h)$ is given by

$$\lambda_h^{-1}(u_h)\varphi_h = Q_h(u_h)\big((I_h\lambda(\mathcal{L}_h u_h))^{-1}\varphi_h\big), \quad \Lambda_h^{-1}(y_h) = \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & \lambda_h^{-1}(u_h) \end{pmatrix}. \tag{2.8}$$

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

$$-\,(\Delta_h\varphi_h \mid \psi_h)_{H_h} = (\varphi_h \mid \psi_h)_{V_h}\,, \qquad \mathrm{A}_h = \begin{pmatrix} 0 & \mathrm{Id} \\ \Delta_h & 0 \end{pmatrix}, \qquad \varphi_h,\psi_h \in V_h. \tag{2.9}$$

Note that Δ_h is symmetric and A_h is skew-symmetric with respect to H_h and X_h, respectively, but they are not uniformly bounded with respect to h. The spatially discrete quasilinear wave equation in first-order formulation then reads

$$\Lambda_h(y_h(t))\partial_t y_h(t) = \mathrm{A}_h y_h(t) + G_h(t,y_h), \qquad t \in [0,T], \tag{2.10}$$

with the initial value $y_h(0) = y_h^0$.

2.1. **Choice of the initial value.** As already mentioned, an appropriately chosen initial value is a key ingredient in the subsequent error analysis. An ideal initial value would include the adjoint lift operator $\mathcal{L}_h^{V*}$ defined in (2.4b). However, in order to compute this operator, integrals over the exact domain Ω have to be evaluated.

We thus propose an alternative that involves to use a finite element space of degree $k' \geq k+1$ denoted by $\widetilde{V}_h$ over the computational domain $\widetilde{\Omega}_h$. Further, let $\widetilde{\mathcal{L}}_h$ and $\widetilde{I}_h$ be the corresponding lift and interpolation operators. Then, for $u \in H^2$, we define the modified Ritz map $\widetilde{R}_h u$ via

$$\Big(\widetilde{R}_h u \mid \varphi_h\Big)_{V_h} = \Big(\widetilde{I}_h u \mid \widetilde{\mathcal{L}_h}^{-1}\mathcal{L}_h\varphi_h\Big)_{\widetilde{V}_h}\,, \qquad \varphi_h \in V_h. \tag{2.11}$$

We use this operator together with the interpolation to define the initial value by

$$y_h^0 = \begin{pmatrix} u_h^0 \\ v_h^0 \end{pmatrix} = \begin{pmatrix} \widetilde{R}_h u^0 \\ I_h v^0 \end{pmatrix}. \tag{2.12}$$

In Appendix A, we prove the following approximation property and discuss the computation of $\widetilde{R}_h$.

Proposition 2.4. *For $u^0 \in H^{k+2}(\Omega) \cap V$, the difference of the adjoint lift $\mathcal{L}_h^{V*}$ defined in (2.4b) and $\widetilde{R}_h$ in (2.11) satisfies the bound*

$$\big\|\mathcal{L}_h^{V*} u^0 - \widetilde{R}_h u^0\big\|_{H^1(\Omega_h)} + h\big\|\Delta_h\big(\mathcal{L}_h^{V*} u^0 - \widetilde{R}_h u^0\big)\big\|_{L^2(\Omega_h)} \le Ch^{k+1} \left\|u^0\right\|_{H^{k+2}},$$

where the constant C is independent of h.

We emphasize that the precise construction of the initial value is not important in the error analysis, but only the bounds obtained in Proposition 2.4. Hence, if we can compute the adjoint lift exactly, which is the Ritz projection in the conforming case, then one can also choose $u_h^0 = \mathcal{L}_h^{V*} u^0$. However, we cannot make the standard choice $u_h^0 = I_h u^0$, since this would imply the statement of Proposition 2.4 only with k instead of $k+1$.

2.2. **Main result for the semi-discretization in space.** Before we state our main error bounds we chose some exponent p^*, depending on the dimension $N = 1, 2, 3$, as

$$N < p^* \begin{cases} \le \infty, & N = 1, \\ < \infty, & N = 2, \\ < 6, & N = 3. \end{cases} \tag{2.13}$$

This choice in particular implies the Sobolev embeddings

$$H^1 \hookrightarrow L^{p^*} \qquad \text{and} \qquad H^2 \hookrightarrow W^{1,p^*} \hookrightarrow L^\infty. \tag{2.14}$$

Our first main result gives an error bound on the spatially discrete solution defined in (2.10), and the proof is given in Section 3. Recall the fractional powers of the Dirichlet Laplacian denoted by $\mathcal{H}_k \coloneqq \mathcal{D}\big((-\Delta)^{k/2}\big)$.

Theorem 2.5. *Let $\partial\Omega \in C^{k+1}$, and Assumption 2.1 hold. Further, let the solution u satisfy*

$$\begin{aligned} u \in C([0,T], \mathcal{H}_3 \cap H^{k+3}(\Omega)) \cap C^2([0,T], V \cap W^{k+1,\infty}(\Omega)), \\ \lambda(u) \in C([0,T], W^{k+1,\infty}(\Omega)), \quad g(\cdot, u, \partial_t u) \in C([0,T], H^{k+1}(\Omega)), \end{aligned} \tag{2.15}$$

and choose the initial value (2.12). Then, there is $h_0 > 0$ such that for all $h \le h_0$, it holds for $t \in [0,T]$

$$\|u(t) - \mathcal{L}_h u_h(t)\|_{W^{1,p^*}(\Omega)} + \|\partial_t u(t) - \mathcal{L}_h v_h(t)\|_{H^1(\Omega)} \le Ch^k$$

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

Using (2.14), the theorem implies convergence in the maximum norm for u_h and in L^{p^*} for v_h, and is in particular applicable to linear finite elements. We note that the results from the literature so far had the limitation $k \ge 2$.

2.3. **Main results for full discretization.** We further discuss the convergence of four different fully discrete schemes. We recall that by $\tau > 0$ we denote the time step size and define for $n = 0, \dots, N$ the times $t^n = n\tau$, with $T = N\tau$. The fully discrete approximations are given by $u_h^n \approx u(t^n)$ and $v_h^n \approx \partial_t u(t^n)$. The proofs of the convergence results are given in Section 4.

Semi-implicit Euler method. For a variant of the implicit Euler method, we introduce the discrete derivative

$$\partial_\tau a_n \coloneqq \tfrac{1}{\tau}\big(a^n - a^{n-1}\big), \quad n \geq 1, \qquad \partial_\tau a_0 \coloneqq a_0, \tag{2.16}$$

and consider as in [10,22] the semi-implicit Euler method

$$\Lambda_h(y_h^n)\partial_\tau y_h^{n+1} = \mathrm{A}_h y_h^{n+1} + G_h(t^n, y_h^n), \quad n \geq 0, \tag{2.17}$$

by freezing the nonlinear parts at the numerical approximation in the last step. The computation of the next approximation thus only requires the solution of a linear system. For the analysis we impose the following weak CFL-type condition

$$\tau \leq c h^{N/p^* + \varepsilon_0} \tag{2.18}$$

with p^* from (2.13) and some arbitrary $\varepsilon_0 > 0$. This yields the following convergence result.

Theorem 2.6. *Let $\partial\Omega \in C^{k+1}$, and Assumption 2.1 hold. Further, let the solution u in addition to (2.15) satisfy*

$$u \in C^3([0,T], L^2(\Omega)),$$

and choose the initial value (2.12). Then, under the condition (2.18) there are $h_0, \tau_0 > 0$ such that for all $h \leq h_0$ and $\tau \leq \tau_0$, it holds for $0 \leq t^n \leq T$

$$\|u(t^n) - \mathcal{L}_h u_h^n\|_{W^{1,p^*}(\Omega)} + \|\partial_t u(t^n) - \mathcal{L}_h v_h^n\|_{H^1(\Omega)} \leq C\big(\tau + h^k\big)$$

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

We emphasize that the CFL-type condition in (2.18) is essentially no restriction for $N = 2$ since p^* can be chosen arbitrarily large due to (2.13). For $N = 3$, the CFL roughly yields $\tau \lesssim h^{1/2+\varepsilon}$. However, even for $k = 1$, the error behaves as $\tau + h$, and one would choose $\tau \sim h$ anyway.

Semi-implicit midpoint rule or Crank–Nicolson scheme. As a second-order in time method, we consider a variant of the midpoint rule proposed in [28]

$$\Lambda_h(\bar{y}_h^{n+1/2})\partial_\tau y_h^{n+1} = \mathrm{A}_h y_h^{n+1/2} + G_h(t^{n+1/2}, \bar{y}_h^{n+1/2}), \quad n \geq 1, \tag{2.19a}$$

with average $y_h^{n+1/2}$ and extrapolation $\bar{y}_h^{n+1/2}$ given by

$$y_h^{n+1/2} = \tfrac{1}{2}\big(y_h^{n+1} + y_h^n\big), \quad \bar{y}_h^{n+1/2} = \tfrac{3}{2}y_h^n - \tfrac{1}{2}y_h^{n-1}. \tag{2.19b}$$

The first approximation y^1 is computed with the Euler method (2.17), and as before, in every time step only a linear system has to be solved. For the analysis of the second-order method, we can weaken the CFL-type condition compared to (2.18) and require only

$$\tau \leq c h^{N/2p^* + \varepsilon_0}. \tag{2.20}$$

Theorem 2.7. *Let $\partial\Omega \in C^{k+1}$, and Assumption 2.1 hold. Further, let the solution u in addition to (2.15) satisfy*

$$u \in C^2([0,T],\mathcal{H}_3) \cap C^3([0,T],\mathcal{H}_2) \cap C^4([0,T],L^2(\Omega)),$$

and choose the initial value (2.12). Then, under the condition (2.20) there are $h_0, \tau_0 > 0$ such that for all $h \leq h_0$ and $\tau \leq \tau_0$, it holds for $0 \leq t^n \leq T$

$$\|u(t^n) - \mathcal{L}_h u_h^n\|_{W^{1,p^*}(\Omega)} + \|\partial_t u(t^n) - \mathcal{L}_h v_h^n\|_{H^1(\Omega)} \leq C(\tau^2 + h^k),$$

where C is independent of h and τ.

Since, there is again essentially no CFL-type condition for $N = 2$, we only discuss the case $N = 3$. We require $\tau \lesssim h^{1/4+\varepsilon}$, whereas in [31] not only $k \geq 2$ but also $\tau \lesssim h^{3/4+\varepsilon}$ has to be imposed.

Exponential Euler method. We turn to exponential methods which employ the variation-of-constants formula and an exact evaluation of the matrix exponential applied to a vector. For the approximation $y_h^n \approx y(t^n)$, we use the shorthand notation $\mathbf{A}_h^n = \Lambda_h^{-1}(y_h^n)\mathrm{A}_h$ and consider the method which was proposed in [12]

$$\begin{aligned} y_h^{n+1} &= \mathrm{e}^{\tau \mathbf{A}_h^n} y_h^n + \tau\varphi_1(\tau\mathbf{A}_h^n) G_h(t^n, y_h^n) \\ &= y_h^n + \tau\varphi_1(\tau\mathbf{A}_h^n)\big(\mathbf{A}_h^n y_h^n + \Lambda_h^{-1}(y_h^n) G_h(t^n, y_h^n)\big) \end{aligned}$$

with the analytic function $\varphi_1(z) = \int_0^1 e^{sz}\, ds$. We obtain the following error bound.

Theorem 2.8. *Let $\partial\Omega \in C^{k+1}$, and Assumption 2.1 hold. Further, let the solution u satisfy (2.15), and choose the initial value (2.12). Then, under the condition (2.18) there are $h_0, \tau_0 > 0$ such that for all $h \leq h_0$ and $\tau \leq \tau_0$, it holds for $0 \leq t^n \leq T$*

$$\|u(t^n) - \mathcal{L}_h u_h^n\|_{W^{1,p^*}(\Omega)} + \|\partial_t u(t^n) - \mathcal{L}_h v_h^n\|_{H^1(\Omega)} \leq C\big(\tau + h^k\big)\,,$$

where C is independent of h and τ.

We note that the CFL-type condition is the same as in the error bound of the semi-implicit Euler in Theorem 2.6.

Exponential midpoint rule. A second-order exponential variant is for example given by the exponential midpoint rule proposed in [12]. Using the notation in (2.19b), we define

$$\mathbf{A}_h^{n+1/2} \coloneqq \Lambda_h^{-1}(\bar{y}_h^{n+1/2})\mathrm{A}_h$$

and consider the scheme

$$\begin{aligned} y_h^{n+1} &= \mathrm{e}^{\tau\mathbf{A}_h^{n+1/2}} y_h^n + \tau\varphi(\tau\mathbf{A}_h^{n+1/2})\Lambda_h^{-1}(\bar{y}_h^{n+1/2}) G_h(t^{n+1/2}, \bar{y}_h^{n+1/2}) \\ &= y_h^n + \tau\varphi(\tau\mathbf{A}_h^{n+1/2})\big(\mathbf{A}_h^{n+1/2} y_h^n + \Lambda_h^{-1}(\bar{y}_h^{n+1/2}) G_h(t^{n+1/2}, \bar{y}_h^{n+1/2})\big). \end{aligned}$$

Employing the techniques established for the proofs of Theorems 2.6 and 2.8, and combining them with the techniques in [12], allow for a convergence result as in Theorem 2.7 under the weaker CFL-type condition (2.20).

Theorem 2.9. *Let $\partial\Omega \in C^{k+1}$, and Assumption 2.1 hold. Further, let the solution u in addition to (2.15) satisfy*

$$\begin{aligned} u &\in C^1([0,T],H^{k+3}) \cap C^3([0,T],W^{k+1,\infty}(\Omega)) \cap C^4([0,T],H^1(\Omega)), \\ \lambda(u) &\in C^3([0,T],W^{k+1,\infty}(\Omega)), \quad g(\cdot,u,\partial_t u) \in C^1([0,T],\mathcal{H}_2) \cap C^3([0,T],H^1(\Omega)), \end{aligned}$$

and choose the initial value (2.12)*. Then, under the condition* (2.20) *there are* $h_0, \tau_0 > 0$ *such that for all* $h \le h_0$ *and* $\tau \le \tau_0$*, it holds for* $0 \le t^n \le T$

$$\|u(t^n) - \mathcal{L}_h u_h^n\|_{W^{1,p^*}(\Omega)} + \|\partial_t u(t^n) - \mathcal{L}_h v_h^n\|_{H^1(\Omega)} \le C(\tau^2 + h^k),$$

where C *is independent of* h *and* τ*.*

2.4. **Numerical experiments.** To illustrate our theoretical findings, we present some numerical experiments for the non-exponential methods. We first illustrate the optimality of our error bounds using a smooth solution, and then consider the formation of a shock wave.

2.4.1. *Smooth solution.* Let $\Omega = B_1(0) \subset \mathbb{R}^2$ be the two-dimensional unit sphere and consider equation (1.1) from Example 2.2 with $\alpha = -\frac{1}{6}$ and data given by

$$u^0(\mathbf{x}) = \frac{1}{10}\sin(\pi r^2)^3 x_1 x_2\,, \qquad v^0(\mathbf{x}) = \frac{1}{10}\sin(\pi r^2)^3 x_1 x_2\,,$$
$$\lambda(u) = 1 - \frac{1}{2}u^2, \qquad g(t,u,v) = uv^2 + \widehat{f}(t),$$

where $r^2 = |\mathbf{x}|^2$. The additional forcing term $\widehat{f}$ is chosen such that the exact solution is given by

$$u(t,\mathbf{x}) = \frac{1}{10}\mathrm{e}^t \sin(\pi r^2)^3 x_1 x_2.$$

A simple calculation shows that the regularity assumptions of Theorems 2.5 to 2.7 are satisfied. The scaling by a factor 10 is used to approximately normalize the $W^{1,\infty}$-norm of the solution u.

Discretization. We discretize in space using the mass and stiffness matrices

$$\big(M_h(u_h)\big)_{i,j} := \Big((I_h^{\Omega_h}\lambda(u_h))\varphi_i \mid \varphi_j\Big)_{L^2(\Omega_h)}, \quad \big(\widetilde{M}_h\big)_{i,j} := (\varphi_i \mid \varphi_j)_{L^2(\Omega_h)}\,,$$
$$\big(L_h\big)_{i,j} := (\nabla\varphi_i \mid \nabla\varphi_j)_{L^2(\Omega_h)}\,,$$

where we denote by $(\varphi_i)_i$ the nodal basis of V_h. Then the discrete solution in (2.10) satisfies

$$M_h\big(u_h(t)\big)\partial_{tt}u_h(t) = -L_h u_h(t) + \widetilde{M}_h I_h^{\Omega_h} g(t, u_h(t), v_h(t)),$$

by abusing the notation for the coefficient vectors and their corresponding function in V_h. The Euler method in (2.17) is then given for $n \ge 0$ by

$$\big(M_h^n + \tau^2 L_h\big)v_h^{n+1} = M_h^n v_h^n - \tau L_h u_h^n + \tau\widetilde{M}_h I_h^{\Omega_h} g(t^n, u_h^n, v_h^n),$$
$$u_h^{n+1} = u_h^n + \tau v_h^{n+1},$$

where we abbreviate $M_h^n = M_h\big(u_h^n\big)$. For the fully discrete midpoint rule, (2.19) is then given for $n \ge 1$ by

$$\begin{aligned}\big(M_h^{n+1/2} + \frac{\tau^2}{4}L_h\big)v_h^{n+1} &= \big(M_h^{n+1/2} - \frac{\tau^2}{4}L_h\big)v_h^n - \tau L_h u_h^n \\ &\qquad + \tau\widetilde{M}_h I_h^{\Omega_h} g(t^{n+1/2}, \bar{u}_h^{n+1/2}, \bar{v}_h^{n+1/2}),\\ u_h^{n+1} &= u_h^n + \frac{\tau}{2}\big(v_h^n + v_h^{n+1}\big),\end{aligned}$$

denoting the extrapolations by $\bar{u}_h^{n+1/2} = \frac{3}{2}u_h^n - \frac{1}{2}u_h^{n-1}$ and $\bar{v}_h^{n+1/2} = \frac{3}{2}v_h^n - \frac{1}{2}v_h^{n-1}$, and the mass matrix by $M_h^{n+1/2} = M_h\big(\bar{u}_h^{n+1/2}\big)$. For the step $n = 0$, we use the Euler scheme from above. We implemented the numerical experiments in `C++` using

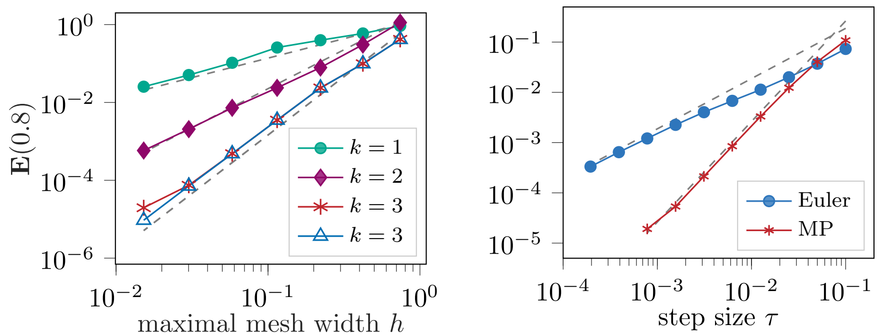


FIGURE 1. Left: error $\mathbf{E}(0.8)$ of the semi-implicit midpoint rule (with time step size $\tau = 8 \cdot 10^{-4}$ and $\tau = 2.67 \cdot 10^{-4}$) combined with finite elements of order $k = 1, 2, 3$ plotted against the mesh width h. The dashed lines indicate order h^k for $k = 1, 2, 3$. Right: Error $\mathbf{E}(0.8)$ of the semi-implicit Euler method and midpoint rule combined with finite elements of order $k = 3$ ($h = 1.52 \cdot 10^{-2}$) plotted against the time step size τ. The dashed lines indicate order 1 and 2.

the finite element library `deal.II` (version 9.4) [2, 6]. A precise description of the implementation can be found for example in [29, Ch. 6.5.1]. For the implementation of the initial value in (2.12), we refer to Appendix A. Concerning the computational costs, let us note that in each step the right-hand side as well as the mass matrix have to be assemble. The stiffness matrix is stored after assembling it before the time stepping. In addition, a linear system for the sparse matrix $M_h + \frac{\tau^2}{4} L_h$ has to be solved in each step using the conjugate gradient method. The codes written by Malik Scheifinger under the authors' supervision to reproduce the experiments are available at `https://doi.org/10.5445/IR/1000163947`.

Numerical results. For the problem described above, we performed experiments for the time and space discretization, where we used finite elements of order $k = 1, 2, 3$. In the error bounds of Section 2, for $N = 2$, the norm $W^{1,p} \times H^1$ is used for $p < \infty$ arbitrarily large. Hence, we chose $p = \infty$ in our experiments, but note that the plots were qualitatively very similar for finite p. Since the computation of the lift of a finite element function is very laborious, and in application usually also not available, 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

$$\mathbf{E}(t) := \|u_h(t) - I_h u(t)\|_{W^{1,\infty}(\Omega_h)} + \|v_h(t) - I_h \partial_t u(t)\|_{H^1(\Omega_h)},$$

for the nodal interpolation operator I_h which is of the same order by the standard interpolation estimates. Note that in practice, one is only interested in u_h, and the computation of the error here is only relevant to confirm our theoretical error bounds.

In the left part of Figure 1, the convergence of the error with respect to the spatial mesh width h is shown when using the semi-implicit midpoint rule with $\tau = 8 \cdot 10^{-4}$. We observe that for finite elements of order k the error converges with order k in space as predicted by Theorems 2.5 and 2.7 until the error for $k = 3$

is dominated by the error of the temporal approximation. For $k = 3$, we ran the same experiment again with the smaller time step size $\tau = 2.67 \cdot 10^{-4}$ to remove this plateau. Running the same experiment with the semi-implicit Euler method instead of the midpoint rule, yields a qualitatively similar picture. Due to slower convergence in time, the error already stagnates at about 10^{-3}.

In the right part of Figure 1, we consider the convergence of the error with respect to the time step size τ for the semi-implicit Euler method and midpoint rule. In space, we discretized with finite elements of order $k = 3$ and $h = 1.52 \cdot 10^{-2}$ such that the spatial error is negligible. Aligning to Theorem 2.6, we observe convergence of order 1 in time for the Euler method and, confirming Theorem 2.7, convergence of order 2 for the midpoint rule.

2.4.2. *Steepening wave.* In this second experiment, we consider the formation of a shock wave which is an often observed phenomenon in nonlinear waves. Since we are in a bounded domain, we force the wave to form a large gradient close to the origin. To this end, we chose our data by

$$u^0(\mathbf{x}) = (1-r^2)^3 \arctan(x_1)\,, \qquad v^0(\mathbf{x}) = -(1-r^2)^3 \frac{\alpha x_1}{x_1^2+1}\,,$$
$$\lambda(u) = 1 - \frac{1}{2}u^2, \qquad g(t,u,v) = uv^2 + \widehat{f}(t),$$

where $r^2 = |\mathbf{x}|^2$. The additional forcing term $\widehat{f}$ is chosen such that the exact solution is given by

$$u(t,\mathbf{x}) = (1-r^2)^3 \arctan\Big(\frac{x_1}{1-\alpha t}\Big). \tag{2.21}$$

We observe that for $\alpha t \to 1$, the maximum norm of ∇u tends to infinity. We thus simulate up to the end time $T = 1$ for different, increasing values of $\alpha < 1$, and in Figure 2 we depicted the corresponding solutions at the end time. The discretization in space and time is performed as described in Section 2.4.1.

Numerical results. We restrict ourselves to the approximation quality in the spatial discretization in this case, and thus apply the semi-implicit midpoint rule with $\tau = 8 \cdot 10^{-4}$ and linear and quadratic finite elements. We then use increasing values of $\alpha = 0.6, 0.8, 0.9, 0.95$, which can be translated into simulating closer to the blow-up point $\alpha t = 1$. We depicted the convergence in Figure 3. In the linear case, we observe that we obtain a reasonable approximation for moderate values of α, which correspond to the smooth case. However, when a shock occurs our method suffers from large errors due to the large gradient. Compared to the linear polynomials, using quadratic polynomials appears to be advantageous, not only because of the better resolution near the blow-up, but also because of smaller error constants. Nevertheless, also in the quadratic case, the error constants become large for $\alpha \to 1$.

2.5. **Additional results for isoparametric finite elements.** In this section, we provide further estimates on the spatially discrete objects which are used throughout the paper.

As shown in [16, Thm. 5.9], we have for the nodal interpolation operator for $m \in \{0,1\}$, $1 \le p \le \infty$, and $1 \le \ell \le k$ the estimates

$$\|(\mathrm{Id} - \mathcal{L}_h I_h^e)\varphi\|_{W^{m,p}(\Omega)} \lesssim h^{\ell+1-m} \|\varphi\|_{W^{\ell+1,p}(\Omega)}\,, \qquad \varphi \in W^{\ell+1,p}(\Omega). \tag{2.22}$$

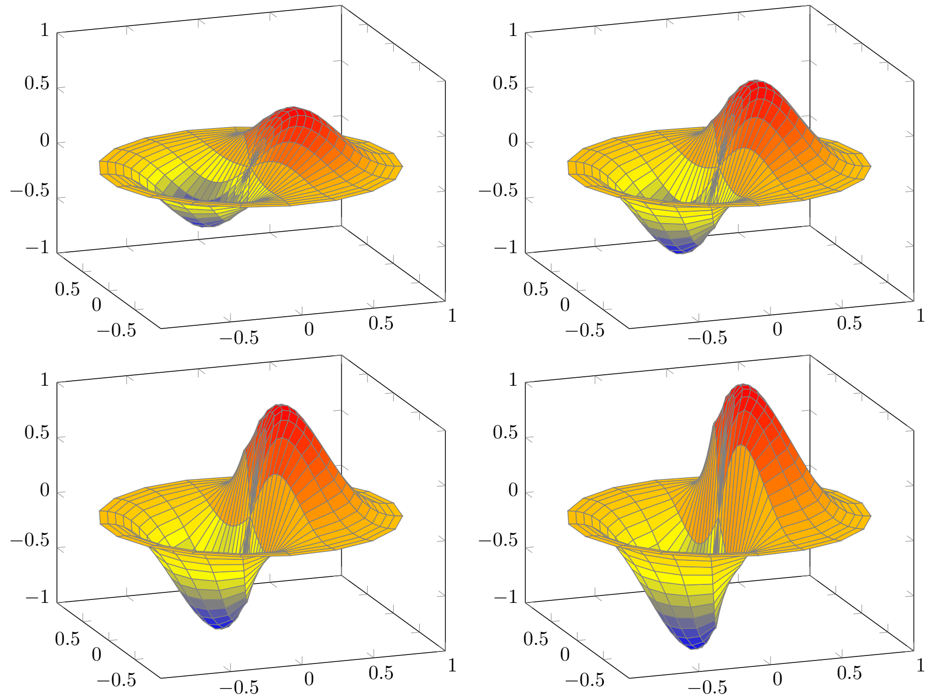


FIGURE 2. Plots of the (exact) shock wave solution given in (2.21) for the final time $t = 1$ and the different values of $\alpha = 0.6, 0.8, 0.9, 0.95$.

Further, by [9, Thm. 3.1.6] $\ell = 0$ is allowed for $N < p \leq \infty$. Another crucial property of the interpolation concerns the stability when applied to the product of functions. We give a proof in Appendix B.

Lemma 2.10. *Let $\psi_h \in V_h$, $\delta > 0$, and $\varphi \in W^{1,N+\delta}(\Omega)$. Then,*

$$\|I_h(\varphi\,\mathcal{L}_h\psi_h)\|_{L^2} \leq C\,\|\varphi\|_{L^\infty}\,\|\psi_h\|_{L^2}\,,$$
$$\|I_h(\varphi\,\mathcal{L}_h\psi_h)\|_{H^1} \leq C\,\|\varphi\|_{W^{1,N+\delta}}\,\|\psi_h\|_{H^1}\,,$$

where the constant $C > 0$ is independent of h.

Concerning the adjoint lifts defined in (2.4), we show in Appendix B the following bounds for $1 \leq \ell \leq k$

$$\left\|\mathcal{L}_h^{H*}\varphi\right\|_{H_h} \lesssim \|\varphi\|_{L^2(\Omega)}\,, \qquad \varphi \in L^2(\Omega), \tag{2.23a}$$
$$\left\|(I_h - \mathcal{L}_h^{H*})\varphi\right\|_{H_h} \lesssim h^{\ell+1}\,\|\varphi\|_{H^{\ell+1}(\Omega)}\,, \qquad \varphi \in H^{\ell+1}(\Omega) \cap V. \tag{2.23b}$$

An interpolation argument between [16, Lem. 3.8] and [14, Thm. 2.5], yields

$$\left\|(\mathrm{Id} - \mathcal{L}_h\mathcal{L}_h^{V*})\varphi\right\|_{W^{1,p}(\Omega)} \lesssim h^{\ell}\,\|\varphi\|_{W^{\ell+1,p}(\Omega)}\,, \qquad \varphi \in H^{\ell+1}(\Omega) \cap V, \tag{2.24}$$

for $2 \leq p \leq \infty$, $0 \leq \ell \leq k$. We will further make use of the inverse estimates, cf. [8, Thm. 4.5.11] or [30, Lem. 5.6],

$$\|\varphi_h\|_{V_h} \leq Ch^{-1}\,\|\varphi_h\|_{L^2(\Omega_h)}\,, \qquad \|\varphi_h\|_{L^q} \leq Ch^{N/q-N/p}\,\|\varphi_h\|_{L^p}\,, \tag{2.25}$$

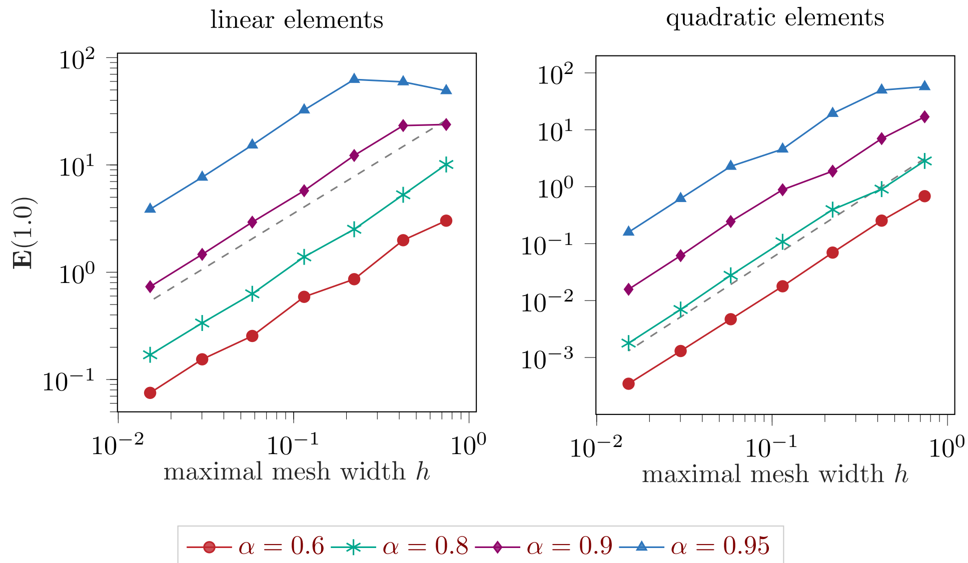


FIGURE 3. Error $\mathbf{E}(1.0)$ of the semi-implicit midpoint rule (with time step size $\tau = 8 \cdot 10^{-4}$) combined with finite elements of order $k = 1$ (left) and order $k = 2$ (right) plotted against the mesh width h for the values of $\alpha = 0.6, 0.8, 0.9, 0.95$. The dashed lines indicate order h or h^2, respectively.

for $1 \le p \le q \le \infty$.

For $u_h \in V_h$ with $\|u_h\|_{L^\infty} \le \widehat{r}_\infty$ and $\|u_h\|_{W^{1,N+\delta}} \le r$, we define an inner product for $\varphi, \psi \in L^2(\Omega_h)$ and the corresponding L^2-projection $Q_h(u_h) \colon L^2(\Omega_h) \to V_h$ used in (2.8) for $\psi_h \in V_h$ by

$$(\varphi \mid \psi)_{\lambda_h} \coloneqq (I_h \lambda(\mathcal{L}_h u_h)\, \varphi \mid \psi)_{L^2(\Omega_h)}\,, \qquad (Q_h(u_h)\psi \mid \psi_h)_{\lambda_h} = (\psi \mid \psi_h)_{\lambda_h}\,,$$

and obtain by the standard techniques for $p \in [2, \infty]$ and $\psi \in L^p, \varphi \in H^1(\Omega_h)$

$$\|\pi_h \psi\|_{L^p(\Omega_h)} \lesssim \|\psi\|_{L^p(\Omega_h)}\,, \quad \|\pi_h \varphi\|_{H^1(\Omega_h)} \lesssim \|\varphi\|_{H^1(\Omega_h)}\,, \tag{2.26a}$$

$$\|Q_h(u_h)\psi\|_{L^p(\Omega_h)} \lesssim \|\psi\|_{L^p(\Omega_h)}\,, \quad \|Q_h(u_h)\varphi\|_{H^1(\Omega_h)} \lesssim \|\varphi\|_{H^1(\Omega_h)}\,, \tag{2.26b}$$

see for example [35] in the case $p = \infty$. The constants are controlled by the norms of u_h in L^∞ and $W^{1,N+\delta}$.

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

$$\boldsymbol{\mathcal{L}_h} = \begin{pmatrix} \mathcal{L}_h & 0 \\ 0 & \mathcal{L}_h \end{pmatrix}, \qquad \boldsymbol{\mathcal{L}_h^*} = \begin{pmatrix} \mathcal{L}_h^{V*} & 0 \\ 0 & \mathcal{L}_h^{H*} \end{pmatrix}, \qquad J_h = \begin{pmatrix} \mathcal{L}_h^{V*} & 0 \\ 0 & \mathcal{L}_h^{V*} \end{pmatrix}, \tag{2.27}$$

which are bounded uniformly in h due to (2.2), (2.23), and (2.24). From the proof of [20, Lem. 4.7], we then obtain the identity

$$\mathrm{A}_h J_h = \boldsymbol{\mathcal{L}_h^*} \mathrm{A}, \tag{2.28}$$

which is used several times in the proofs.

3. Error analysis for the space discretization

In this section, we give the proof of Theorem 2.5. We decompose the error into

$$\begin{aligned} y(t) - \boldsymbol{\mathcal{L}_h} y_h(t) &= \big(\mathrm{Id} - \boldsymbol{\mathcal{L}_h} J_h\big) y(t) + \boldsymbol{\mathcal{L}_h}\big(J_h y(t) - y_h(t)\big) \\ &=: e_{J_h}(t) + \boldsymbol{\mathcal{L}_h} e_h(t), \end{aligned}$$

where the projection error e_{J_h} is easily bounded using (2.24). The first part of the proof consists in reducing the bound on $\|e_h\|_{W^{1,p^*}\times H^1}$ to an estimate in the stronger norm induced by $\|\mathrm{A}_h \cdot\|_{X_h}$, and not in the standard X_h-norm.

The second part consists in establishing the stronger norm bound on $\|\mathrm{A}_h e_h\|_{X_h}$ in Section 3.2. We note that a key idea is to set up an appropriate solution space for the numerical approximation, see (3.3) below, which allows for an appropriate formulation of the error equation. We give a detailed explanation in Remark 3.6.

3.1. Reduction to stronger norm estimates.

For p^* defined in (2.13), we chose some fixed $\delta > 0$ such that

$$\frac{1}{2} - \frac{1}{N+\delta} \le \frac{1}{p^*}, \tag{3.1}$$

a radius $r_\infty < \widehat{r}_\infty$ from Assumption 2.1, and another radius $r'_\infty > 0$, such that

$$\|u\|_{L^\infty(L^\infty)} \le r_\infty, \quad \text{and} \quad \max\big\{\|u\|_{L^\infty(W^{1,N+\delta})}, \|\partial_t u\|_{L^\infty(W^{1,N+\delta})}\big\} \le \tfrac{1}{2} r'_\infty, \tag{3.2}$$

where $\|x\|_{L^\infty(X)} := \max_{[0,T]} \|x(t)\|_X$. We denote by t_h^* the time with

$$\begin{aligned} t_h^* := \sup\{t \in [0,T] \mid &\sup_{s\in[0,t]} \|\mathcal{L}_h u_h(s)\|_{L^\infty} \le \widehat{r}_\infty \quad \text{and} \\ &\sup_{s\in[0,t]} \|\mathcal{L}_h u_h(s)\|_{W^{1,N+\delta}}, \sup_{s\in[0,t]} \|\mathcal{L}_h v_h(s)\|_{W^{1,N+\delta}} \le r'_\infty\}. \end{aligned} \tag{3.3}$$

We assume for a moment that the set is not empty and hence $t_h^* > 0$, see Proposition 3.5. The following result is a direct consequence of Lemma 2.10 and the key ingredient to ensure wellposedness of the discrete equation. In addition, it enables us to employ energy techniques in the error analysis.

Lemma 3.1. *Let Assumption 2.1 hold. We have for $t \in [0, t_h^*]$, $1 \le p \le \infty$, and $j = 0, 1$ the bounds*

$$\big\|\partial_t^j \lambda_h(u_h(t))\varphi_h\big\|_{L^p} \le C_\lambda \|\varphi_h\|_{L^p}, \qquad \big\|\partial_t^j \lambda_h^{-1}(u_h(t))\varphi_h\big\|_{L^p} \le C_\lambda \|\varphi_h\|_{L^p},$$

$$\big\|\lambda_h(u_h(t))\varphi_h\big\|_{H^1} \le C_\lambda \|\varphi_h\|_{H^1}, \qquad \big\|\lambda_h^{-1}(u_h(t))\varphi_h\big\|_{H^1} \le C_\lambda \|\varphi_h\|_{H^1},$$

with a constant $C_\lambda > 0$ depending only on λ, its derivatives and $\widehat{r}_\infty, r'_\infty$, but is independent of h and t_h^*.

Proof. We use the definition of λ_h and λ_h^{-1} in (2.6) and (2.8), respectively, to conclude the assertion from Assumption 2.1, the stability in (2.26), the interpolation property (2.22), and (3.3). □

Making extensive use of Lemma 3.1, we show via energy techniques in Section 3.2 the following error bound on

$$\Big(\big\|\Delta_h\big(\mathcal{L}_h^{V*} u(t) - u_h(t)\big)\big\|_{L^2}^2 + \big\|\mathcal{L}_h^{V*}\partial_t u(t) - v_h(t)\big\|_{H^1}^2\Big)^{1/2} = \big\|\mathrm{A}_h e_h(t)\big\|_{X_h}. \tag{3.4}$$

Note that initially the result is only valid as long as the bounds in (3.3) hold.

Proposition 3.2. *Under the assumptions of Theorem 2.5, it holds for* $0 \le t \le t_h^*$

$$\left\| \mathrm{A}_h e_h(t) \right\|_{X_h} \le C h^k,$$

where C *is independent of* h *and* t_h^*.

From this bound, we are able to extract convergence as well as to extend the final time t_h^* beyond T for sufficiently small h. Concerning u_h, we show in the following lemma how to obtain convergence in the maximum norm and first-order Sobolev norms, but postpone the proof to Appendix C. Further, we may directly deduce the bounds on u_h in (3.3). Note that this lemma can be seen as a discrete analogue to (2.14), and is an improved variant of the results in [7, 13]. Similar bounds were already shown in [18, Thm. 1.12] and [36, Thm. 3].

Lemma 3.3. *Let* p^* *be given by* (2.13). *Then, there is a constant* C *independent of* h *such that*

$$\|\varphi_h\|_{L^\infty(\Omega_h)} + \|\varphi_h\|_{W^{1,p^*}(\Omega_h)} \le C \|\Delta_h \varphi_h\|_{L^2(\Omega_h)}$$

for all $\varphi_h \in V_h$. *In the case* $N = 3$, *the statement also holds for* $p^* = 6$.

A further ingredient in the proof of the main result is to employ the H^1-bound on v_h in Proposition 3.2 to derive boundedness in L^∞ and $W^{1,N+\delta}$, and thus extend the final time t_h^*.

Lemma 3.4. *Let* $\varphi_h \in V_h$ *and* $\varphi \in W^{k+1,\infty}(\Omega) \cap V$, *and assume that*

$$\left\| \mathcal{L}_h^{V*} \varphi - \varphi_h \right\|_{H^1(\Omega_h)} \le C h^k. \tag{3.5}$$

Then, we have for p^* *defined in* (2.13) *and* δ *chosen in* (3.1)

$$\|\mathcal{L}_h \varphi_h\|_{L^\infty(\Omega)} \le \|\varphi\|_{L^\infty(\Omega)} + C h^{k-N/p^*},$$

$$\|\mathcal{L}_h \varphi_h\|_{W^{1,N+\delta}(\Omega)} \le \|\varphi\|_{W^{1,N+\delta}(\Omega)} + C h^{k-N/p^*},$$

with a constant C *independent of* h.

Since $k \ge 1$, the choice in (2.13) enables to us to deduce the desired bounds (3.3) in L^∞ and $W^{1,N+\delta}$ from approximation properties in H^1, and hence allows us to extend the final time t_h^*.

Proof of Lemma 3.4. For $\psi_h \in V_h$ we combine the inverse estimate (2.25) and the Sobolev embedding $H^1(\Omega_h) \hookrightarrow L^{p^*}(\Omega_h)$, and conclude by (3.1)

$$\|\psi_h\|_{L^\infty(\Omega_h)} \le C h^{-N/p^*} \|\psi_h\|_{L^{p^*}(\Omega_h)} \le C h^{-N/p^*} \|\psi_h\|_{H^1(\Omega_h)}, \tag{3.6a}$$

$$\|\psi_h\|_{W^{1,N+\delta}(\Omega_h)} \le C h^{N/(N+\delta)-N/2} \|\psi_h\|_{H^1(\Omega_h)} \le C h^{-N/p^*} \|\psi_h\|_{H^1(\Omega_h)}, \tag{3.6b}$$

with a constant C independent of h. For the desired bound, we expand with the adjoint lift $\mathcal{L}_h^{V*}$ and obtain by (2.24)

$$\begin{aligned}
\|\mathcal{L}_h \varphi_h\|_{L^\infty(\Omega)} &\le \|\varphi\|_{L^\infty(\Omega)} + \left\|\varphi - \mathcal{L}_h \mathcal{L}_h^{V*} \varphi\right\|_{L^\infty(\Omega)} + \left\|\mathcal{L}_h \mathcal{L}_h^{V*} \varphi - \mathcal{L}_h \varphi_h\right\|_{L^\infty(\Omega)} \\
&\le \|\varphi\|_{L^\infty(\Omega)} + C h^k \|\varphi\|_{W^{k+1,\infty}(\Omega)} + C \left\|\mathcal{L}_h^{V*} \varphi - \varphi_h\right\|_{L^\infty(\Omega_h)}.
\end{aligned}$$

Since $\mathcal{L}_h^{V*} \varphi - \varphi_h \in V_h$, the first assertion then follows from (3.6a) together with (3.5). The second estimate is derived fully analogously. □

Hence, once we have shown Proposition 3.2, we can give the proof of our first main result.

Proof of Theorem 2.5. Inserting the adjoint lift, we obtain for $t \in [0, t_h^*]$ with (2.24), (3.4), and Lemma 3.3 the bound

$$\|u(t) - \mathcal{L}_h u_h(t)\|_{W^{1,p^*}} \leq \left\|(\mathrm{Id} - \mathcal{L}_h \mathcal{L}_h^{V*})u(t)\right\|_{W^{1,p^*}} + C\,\|\mathrm{A}_h e_h(t)\|_{X_h} \leq Ch^k\,,$$

and similarly

$$\|\partial_t u(t) - \mathcal{L}_h v_h(t)\|_{H^1} \leq \left\|(\mathrm{Id} - \mathcal{L}_h \mathcal{L}_h^{V*})\partial_t u(t)\right\|_{H^1} + C\,\|\mathrm{A}_h e_h(t)\|_{X_h} \leq Ch^k\,,$$

with a constant C independent of h and t_h^*. Hence, it remains to show $t_h^* = T$. Combining the bounds in Proposition 3.2 and Lemma 3.3, we show by (3.2) for h sufficiently small that

$$\|\mathcal{L}_h u_h(t_h^*)\|_{L^\infty(\Omega)} \leq \|u(t_h^*)\|_{L^\infty(\Omega)} + Ch^k < \widehat{r}_\infty\,,$$
$$\|\mathcal{L}_h u_h(t_h^*)\|_{W^{1,N+\delta}(\Omega)} \leq \|u(t_h^*)\|_{W^{1,N+\delta}(\Omega)} + Ch^k < r'_\infty\,,$$

as well as with Proposition 3.2 and Lemma 3.4

$$\|\mathcal{L}_h v_h(t_h^*)\|_{W^{1,N+\delta}(\Omega)} \leq \|\partial_t u(t_h^*)\|_{W^{1,N+\delta}(\Omega)} + Ch^{k-N/p^*} < r'_\infty.$$

Thus, the continuity of the discrete solution y_h and the equivalence of all norms in finite dimensional spaces yields $t_h^* \geq T$. In particular, the statement of Theorem 2.5 is true for $t \in [0, T]$. □

3.2. **Proof of Proposition 3.2.** The rest of this section is devoted to the proof of Proposition 3.2. The first step is to show that the set defined in (3.3) is not empty.

Proposition 3.5. *The initial error satisfies*

$$\|\mathrm{A}_h e_h(0)\|_{X_h} \leq Ch^k,$$

where C is independent of h. In particular, it holds $0 < t_h^* \leq T$.

Proof. The bound directly follows from the choice (2.12), the interpolation properties in (2.22), and the bounds in Proposition 2.4. To show that $t_h^* > 0$, we proceed as in the proof of Theorem 2.5 with $t = 0$ instead of $t = t_h^*$. □

With the aid of Lemma 3.1, we are able to define with $\Lambda_h(y_h)$ from (2.7) the state-dependent inner products

$$(\varphi_h \mid \psi_h)_{\Lambda_h,t} := (\Lambda_h(y_h(t))\varphi_h \mid \psi_h)_{X_h}\,, \qquad t \in [0, t_h^*],\ \varphi_h, \psi_h \in X_h.$$

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

$$c_{\Lambda_h}\,\|\varphi_h\|_{X_h} \leq \|\varphi_h\|_{\Lambda_h,t} \leq C_{\Lambda_h}\,\|\varphi_h\|_{X_h}\,, \qquad t \in [0, t_h^*],\ \varphi_h \in X_h, \tag{3.7}$$

with the constants from Lemma 3.1.

Error equation. We study the bound on the discrete error e_h and derive an evolution equation for it. Inserting the projected solution $J_h y$ of (2.1) in (2.10), we obtain

$$\begin{aligned}\Lambda_h(y_h(t))J_h\partial_t y(t) = \mathrm{A}_h J_h y(t) + G_h(t, J_h y) + \big(\Lambda_h(y_h(t)) - \Lambda_h(J_h y(t))\big)J_h \partial_t y(t)\\ + \delta_h(t)\end{aligned}$$

with defect

$$\begin{aligned}\delta_h(t) = &\big(\Lambda_h(J_h y(t))J_h - J_h\Lambda(y(t))\big)\partial_t y(t)\\ &+ \big(J_h\mathrm{A} - \mathrm{A}_h J_h\big)y(t) + \big(J_h G(t, y) - G_h(t, J_h y)\big).\end{aligned} \tag{3.8}$$

This leads us to the error equation

$$\Lambda_h(y_h(t))\partial_t e_h(t) = \mathrm{A}_h e_h(t) + \Gamma_h(t) + \delta_h(t), \tag{3.9}$$

where the stability term is given by

$$\Gamma_h(t) \coloneqq \big(G_h(t, J_h y(t)) - G_h(t, y_h(t))\big) + \big(\Lambda_h(y_h(t)) - \Lambda_h(J_h y(t))\big) J_h \partial_t y(t). \tag{3.10}$$

Remark 3.6. Let us explain the main differences to the error analysis presented by Maier and Hochbruck [21, 31] and Makridakis [32]. In [21, 31], instead of $\Lambda_h(y_h)$ they use $\Lambda_h(I_h y)$ which has the properties from Lemma 3.1. However, to bound the stability term, inverse inequalities are used which induce restrictions on the polynomial degree and also the CFL-type condition. Our technique is more related to [32], where bounds on $\|u_h\|_{W^{1,\infty}}$ replace (3.3).

However, in both approaches the error analysis is performed in $H^1 \times L^2$. They thus have to impose stronger CFL-type conditions to close the argument. ♢

We introduce the state-dependent operator

$$\mathbf{A}_h(t) = \Lambda_h^{-1}(y_h(t))\mathrm{A}_h$$

and define the modified error as

$$\widetilde{e}_h(t) \coloneqq \mathbf{A}_h(t) e_h(t).$$

Differentiating the term $\Lambda_h(y_h(t))\widetilde{e}_h(t)$ and using (3.9), leads to the following modified error equation

$$\begin{aligned}\Lambda_h(y_h(t))\partial_t \widetilde{e}_h(t) = \mathrm{A}_h \widetilde{e}_h(t) &- \big(\partial_t \Lambda_h(y_h(t))\big)\widetilde{e}_h(t) \\ &+ \mathrm{A}_h \Lambda_h^{-1}(y_h(t))\big(\Gamma_h(t) + \delta_h(t)\big).\end{aligned} \tag{3.11}$$

We state two results on the stability term and the defect, and postpone their proofs to Section 5.

Lemma 3.7. *For $0 \le t \le t_h^*$ it holds*

$$\|\mathrm{A}_h \Gamma_h(t)\|_{X_h} \le C \|\mathrm{A}_h e_h(t)\|_{X_h}$$

with a constant C independent of h and t_h^*.

Similarly, we show the optimal error bound of the defect in the stronger norm.

Lemma 3.8. *For $0 \le t \le t_h^*$ it holds*

$$\|\mathrm{A}_h \delta_h(t)\|_{X_h} \le C h^k$$

with a constant C independent of h and t_h^*.

In addition, we note that by (2.7) and Lemma 3.1 there is a constant C independent of h and t_h^* such that for all $x_h \in X_h$ it holds

$$\|\mathrm{A}_h \Lambda_h(y_h(t)) x_h\|_{X_h} \le C \|\mathrm{A}_h x_h\|_{X_h}, \quad 0 \le t \le t_h^*. \tag{3.12}$$

With these two lemmas and the bound on the initial error in Proposition 3.5, we conclude the remaining estimate.

Proof of Proposition 3.2. We first compute

$$\partial_t \|\widetilde{e}_h(t)\|_{\Lambda_h,t}^2 = \Big(\big(\partial_t \Lambda_h(y_h(t))\big)\widetilde{e}_h(t) \mid \widetilde{e}_h(t)\Big)_{X_h} + 2\left(\Lambda_h(y_h(t))\partial_t \widetilde{e}_h(t) \mid \widetilde{e}_h(t)\right)_{X_h}.$$

Inserting the error equation (3.11), we use the skew-symmetry of A_h and combine the bounds in (3.12) and Lemmas 3.1, 3.7, and 3.8 to obtain

$$\partial_t \|\widetilde{e}_h(t)\|^2_{\Lambda_h,t} \leq C \|\widetilde{e}_h(t)\|^2_{\Lambda_h,t} + Ch^{2k}.$$

The application of a Gronwall lemma together with Proposition 3.5 and (3.7) then yields the assertion. □

4. Error analysis for the full discretization

We carry over the results and techniques established in the last section to the fully discrete schemes. We work with a discrete analogous of (3.3) given by a final time step n^* which allows us to perform the next time step to t^{n^*+1}, that is

$$\begin{aligned} n^* := \max\Big\{ 0 \leq n \leq N-1 \mid \max_{k=0,\dots,n} \big\|\mathcal{L}_h u_h^k\big\|_{L^\infty} \leq \widehat{r}_\infty, \text{ and} \\ \max_{k=0,\dots,n} \max\big\{ \big\|\mathcal{L}_h u_h^k\big\|_{W^{1,N+\delta}}, \big\|\mathcal{L}_h v_h^k\big\|_{W^{1,N+\delta}}, \big\|\mathcal{L}_h \partial_\tau u_h^k\big\|_{W^{1,N+\delta}} \big\} \leq r'_\infty \Big\}. \end{aligned} \tag{4.1}$$

In particular, we will establish $n^* \geq N-1$. Note that by (2.16) formally, we have to show that $n^* \geq 1$ for the last term in (4.1), which can be interpreted as providing both the base cases $n = 0, 1$ in the induction. However, the case $n = 0$ is already covered by Proposition 3.2, such that the set in (4.1) is not empty, and it holds $n^* \geq 0$.

Further, note that similar to Lemma 3.1 we conclude from the bounds in (4.1) that for $0 \leq n \leq n^*$, $1 \leq p \leq \infty$, and $j = 0, 1$ it holds

$$\big\|\partial_\tau^j \lambda_h(u_h^n)\varphi_h\big\|_{L^p} \leq C_\lambda \|\varphi_h\|_{L^p}, \qquad \big\|\partial_\tau^j \lambda_h^{-1}(u_h^n)\varphi_h\big\|_{L^p} \leq C_\lambda \|\varphi_h\|_{L^p}, \tag{4.2}$$

and the bounds in Lemma 3.1 in the H^1-norm remain valid.

Throughout this section, we employ several times the estimate from Lemma 3.3, and also a straightforward extension of Lemma 3.4 including the temporal convergence rate.

Lemma 4.1. *Let $\varphi_h \in V_h$ and $\varphi \in W^{k+1,\infty}(\Omega) \cap V$, and assume that for some $\ell \in \{1,2\}$ it holds*

$$\big\|\mathcal{L}_h^{V*}\varphi - \varphi_h\big\|_{H^1(\Omega_h)} \leq C\big(\tau^\ell + h^k\big).$$

Then, we have

$$\begin{aligned} \|\mathcal{L}_h\varphi_h\|_{L^\infty(\Omega)} &\leq \|\varphi\|_{L^\infty(\Omega)} + Ch^{-N/p^*}\big(\tau^\ell + h^k\big), \\ \|\mathcal{L}_h\varphi_h\|_{W^{1,N+\delta}(\Omega)} &\leq \|\varphi\|_{W^{1,N+\delta}(\Omega)} + Ch^{-N/p^*}\big(\tau^\ell + h^k\big), \end{aligned}$$

with a constant C independent of h and τ.

4.1. **Euler.** First note that for the Euler method (2.17), we have by construction $\partial_\tau u_h^k = v_h^k$ such that it is sufficient to check the first three conditions. As above, we define the discrete error by $e_h^n = J_h y(t^n) - y_h^n$ and aim to show as in Proposition 3.2 the following bound.

Proposition 4.2. *Under the assumptions of Theorem 2.6, for $0 \leq n \leq n^* + 1$ it holds the bound*

$$\big\|\mathrm{A}_h e_h^n\big\|_{X_h} \leq C\big(\tau + h^k\big),$$

where C is independent of h, τ and n^*.

As in the spatially discrete case, this estimate allows us to immediately conclude our main result.

Proof of Theorem 2.6. We proceed along the lines of the proof of Theorem 2.5 to conclude the convergence up to t^{n^*+1}. In addition, Lemma 3.3 and the CFL-type condition (2.18) together with Lemma 4.1 for $\ell = 1$ further allow us to prove $n^* \geq N-1$ for h, τ sufficiently small, and the assertion is shown for all n. $\square$

The rest of this section is devoted to the proof of Proposition 4.2. In order to derive the error equation, we insert the projected exact solution $J_h y$ of (2.1) in the scheme (2.17) and derive

$$\begin{aligned}\Lambda_h(y_h^n) J_h \partial_\tau y(t^{n+1}) &= \mathrm{A}_h J_h y(t^{n+1}) + G_h(t^n, J_h y(t^n)) \\ &\quad + \big(\Lambda_h(y_h^n) - \Lambda_h(J_h y(t^n))\big) J_h \partial_\tau y(t^{n+1}) + \delta_{\mathrm{Eu}}^{n+1}\end{aligned}$$

with defect $\delta_{\mathrm{Eu}}^{n+1} = \delta_{h,\mathrm{Eu}}^{n+1} + \delta_{\tau,\mathrm{Eu}}^{n+1}$ given by

$$\begin{aligned}\delta_{h,\mathrm{Eu}}^{n+1} &= \big(J_h \mathrm{A} - \mathrm{A}_h J_h\big) y(t^{n+1}) + J_h G(t^n, y(t^n)) - G_h(t^n, J_h y(t^n)) \\ &\quad + \big(\Lambda_h(J_h y(t^n)) J_h - J_h \Lambda(y(t^n))\big) \partial_\tau y(t^{n+1}),\end{aligned} \tag{4.3a}$$

$$\begin{aligned}\delta_{\tau,\mathrm{Eu}}^{n+1} &= J_h \Lambda(y(t^n)) \partial_\tau y(t^{n+1}) - J_h \Lambda(y(t^{n+1})) \partial_t y(t^{n+1}) \\ &\quad + J_h G(t^{n+1}, y(t^{n+1})) - J_h G(t^n, y(t^n)).\end{aligned} \tag{4.3b}$$

This yields the discrete error equation

$$\Lambda_h(y_h^n) \partial_\tau e_h^{n+1} = \mathrm{A}_h e_h^{n+1} + \Gamma_h^n + \delta_{\mathrm{Eu}}^{n+1}, \tag{4.4}$$

where the stability term is given by

$$\Gamma_h^n \coloneqq \big(G_h(t^n, J_h y(t^n)) - G_h(t^n, y_h)\big) + \big(\Lambda_h(y_h^n) - \Lambda_h(J_h y(t^n))\big) J_h \partial_\tau y(t^{n+1}). \tag{4.5}$$

In order to obtain a recursion for e_h^{n+1}, we recall the state-dependent operator and define the corresponding resolvent

$$\mathbf{A}_h^n = \Lambda_h^{-1}(y_h^n) \mathrm{A}_h, \qquad \mathbf{R}_{\mathrm{Eu},n} \coloneqq \big(I - \tau \mathbf{A}_h^n\big)^{-1}. \tag{4.6}$$

A simple calculation shows that for the inner product

$$(\varphi_h \mid \psi_h)_n \coloneqq (\Lambda_h(y_h^n) \varphi_h \mid \psi_h)_{X_h}, \qquad \varphi_h, \psi_h \in X_h,$$

which satisfies by (4.1) the same bounds as in (3.7), we obtain

$$\|\mathbf{R}_{\mathrm{Eu},n} \varphi_h\|_n \leq \|\varphi_h\|_n,$$

and rewrite (4.4) as

$$e_h^{n+1} = \mathbf{R}_{\mathrm{Eu},n} e_h^n + \tau \mathbf{R}_{\mathrm{Eu},n} \Lambda_h^{-1}(y_h^n) \big(\Gamma_h^n + \delta_{\mathrm{Eu}}^{n+1}\big).$$

Since $\mathbf{A}_h^n$ commutes with $\mathbf{R}_{\mathrm{Eu},n}$, we obtain

$$\mathbf{A}_h^n e_h^{n+1} = \mathbf{R}_{\mathrm{Eu},n} \mathbf{A}_h^n e_h^n + \tau \mathbf{R}_{\mathrm{Eu},n} \mathbf{A}_h^n \Lambda_h^{-1}(y_h^n) \big(\Gamma_h^n + \delta_{\mathrm{Eu}}^{n+1}\big)$$

which has to be resolved. Proceeding as in Lemma 3.7, and noting that for any norm it holds

$$\|\partial_\tau y(t^n)\| \leq \max_{t \in [t^{n-1}, t^n]} \|\partial_t y(t)\|, \tag{4.7}$$

we have for $0 \leq n \leq n^*$ the stability bound

$$\|\mathrm{A}_h \Gamma_h^n\|_{X_h} \leq C \|\mathrm{A}_h e_h^n\|_{X_h}$$

with a constant C independent of h, τ and n^*. Similarly, we show the optimal consistency error of the defect in the stronger norm, see Section 5 for the proof.

Lemma 4.3. *For $0 \le n \le n^*$ it holds*

$$\left\| \mathrm{A}_h \delta_{Eu}^{n+1} \right\|_{X_h} \le C\big(\tau + h^k\big)$$

with a constant C independent of h, τ and n^*.

Hence, we have already established the estimate

$$\left\| \mathbf{A}_h^n e_h^{n+1} \right\|_n \le \left\| \mathbf{A}_h^n e_h^n \right\|_n + C\tau \left\| \mathrm{A}_h e_h^n \right\|_{X_h} + C\tau\big(\tau + h^k\big), \tag{4.8}$$

and the last step towards the main result is to change the norms.

Lemma 4.4. *For $1 \le n \le n^*$ it holds for all $\varphi_h \in V_h$*

$$\left\| \mathbf{A}_h^n \varphi_h \right\|_n \le (1 + C\tau) \left\| \mathbf{A}_h^{n-1} \varphi_h \right\|_{n-1}$$

with a constant C independent of h, τ and n^*.

Proof. Expanding the norm as

$$\begin{aligned} \left\| \mathbf{A}_h^n \varphi_h \right\|_n^2 &= \big(\mathrm{A}_h \varphi_h \mid \Lambda_h^{-1}(y_h^n) \mathrm{A}_h \varphi_h \big)_{X_h} \\ &= \left\| \mathbf{A}_h^{n-1} \varphi_h \right\|_{n-1}^2 + \tau \left(\mathrm{A}_h \varphi_h \mid \partial_\tau \Lambda_h^{-1}(y_h^n) \mathrm{A}_h \varphi_h \right)_{X_h} \end{aligned}$$

and using (4.2) several times, gives the assertion. □

With this we are able to proof the estimate on $\mathrm{A}_h e_h^{n+1}$.

Proof of Proposition 4.2. We first consider the case $n^* = 0$. Hence, (4.8) with $n = 0$ directly yields the assertion without the use of Lemma 4.4 and hence without any bound on $\partial_\tau u_h^k$. With this, we established $n^* \ge 1$.

In the case $n^* \ge 1$, we employ Lemma 4.4 in (4.8) and make use of the norm equivalences to obtain

$$\left\| \mathbf{A}_h^n e_h^{n+1} \right\|_n \le (1 + C\tau) \left\| \mathbf{A}_h^{n-1} e_h^n \right\|_{n-1} + C\tau\big(\tau + h^k\big).$$

Resolving the recursion and using Proposition 3.5 yields the result. □

4.2. **Midpoint.** The proof is very similar to the Euler method and hence, we only sketch the relevant details. First note that by construction in (2.19) it holds

$$\partial_\tau u_h^k = \tfrac{1}{2}(v_h^k + v_h^{k-1}),$$

such that the last bound in (4.1) does not have to be shown separately. Again, we aim at the following bound.

Proposition 4.5. *Under the assumptions of Theorem 2.7, for $0 \le n \le n^* + 1$ it holds the bound*

$$\left\| \mathrm{A}_h e_h^n \right\|_{X_h} \le C\big(\tau^2 + h^k\big),$$

where C is independent of h, τ and n^*.

Combining Lemma 4.1 with the weaker CFL-type condition (2.20) yields the convergence result.

Proof of Theorem 2.7. As in the proof of Theorem 2.6, the convergence follows directly. To show that $n^* \ge N - 1$, we employ Lemma 4.1 with $\ell = 2$ together with the CFL-type condition (2.20). □

Hence, it remains to show Proposition 4.5. As for the Euler method, we derive the following error equation

$$\Lambda_h(\bar{y}_h^{n+1/2})\partial_\tau e_h^{n+1} = \mathrm{A}_h e_h^{n+1/2} + \Gamma_h^n + \delta_{\mathrm{M}}^{n+1}, \tag{4.9}$$

with a stability term similar to the one in (4.5) satisfying

$$\|\mathrm{A}_h\Gamma_h^n\|_{X_h} \le C\big(\|\mathrm{A}_h e_h^n\|_{X_h} + \big\|\mathrm{A}_h e_h^{n-1}\big\|_{X_h}\big), \tag{4.10}$$

and a composed defect $\delta_{\mathrm{M}}^{n+1} = \delta_{h,\mathrm{M}}^{n+1} + \delta_{\tau,\mathrm{M}}^{n+1}$. The first component is basically the same as $\delta_{h,\mathrm{Eu}}^{n+1}$ in (4.3b), and the second satisfies

$$\begin{aligned} \delta_{\tau,\mathrm{M}}^{n+1} &= J_h\Lambda(\bar{y}^{n+1/2})\partial_\tau y(t^{n+1}) - J_h\Lambda(y(t^{n+1/2}))\partial_t y(t^{n+1/2}) \\ &\quad + J_h\mathrm{A}\big(y(t^{n+1/2}) - \tfrac{1}{2}(y(t^{n+1}) + y(t^n))\big), \\ &\quad + J_h G(t^{n+1/2}, y(t^{n+1/2})) - J_h G(t^{n+1/2}, \bar{y}^{n+1/2}), \end{aligned} \tag{4.11}$$

such that we derive in Section 5 the desired order of convergence.

Lemma 4.6. *For $0 \le n \le n^*$ it holds*

$$\big\|\mathrm{A}_h\delta_M^{n+1}\big\|_{X_h} \le C\big(\tau^2 + h^k\big)$$

with a constant C independent of h, τ and n^*.

We solve for e_h^{n+1} in the error equation (4.9) and define for the solution-dependent operator $\mathbf{A}_h^{n+1/2} = \Lambda_h^{-1}(\bar{y}^{n+1/2})\mathrm{A}_h$ the maps

$$\mathcal{R}_{\pm,n+1/2} \coloneqq I \pm \tfrac{\tau}{2}\mathbf{A}_h^{n+1/2}, \qquad \mathbf{R}_{\mathrm{m},n+1/2} \coloneqq \mathcal{R}_{-,n+1/2}^{-1}\mathcal{R}_{+,n+1/2}.$$

A simple calculation shows that for the inner product

$$(\varphi_n \mid \psi_h)_{n+1/2} \coloneqq \big(\Lambda_h(\bar{y}^{n+1/2})\varphi_n \mid \psi_h\big)_{X_h}, \qquad \varphi_n, \psi_h \in X_h,$$

we have

$$\big\|\mathcal{R}_{-,n+1/2}^{-1}\varphi_h\big\|_{n+1/2} \le \|\varphi_h\|_{n+1/2}, \qquad \big\|\mathbf{R}_{\mathrm{m},n+1/2}\varphi_h\big\|_{n+1/2} = \|\varphi_h\|_{n+1/2}.$$

Rewriting (4.9) and multiplying by $\mathbf{A}_h^{n+1/2}$, we obtain

$$\begin{aligned} \mathbf{A}_h^{n+1/2}e_h^{n+1} &= \mathbf{R}_{\mathrm{m},n+1/2}\mathbf{A}_h^{n+1/2}e_h^n \\ &\qquad + \tau\mathcal{R}_{-,n+1/2}^{-1}\mathbf{A}_h^{n+1/2}\Lambda_h^{-1}(\bar{y}^{n+1/2})\big(\Gamma_h^n + \delta_{\mathrm{M}}^{n+1}\big). \end{aligned} \tag{4.12}$$

Finally, we have as in Lemma 4.4 the following bound when changing the norm.

Lemma 4.7. *For $1 \le n \le n^*$ it holds*

$$\big\|\mathbf{A}_h^{n+1/2}\varphi_h\big\|_{n+1/2} \le (1 + C\tau)\big\|\mathbf{A}_h^{n-1/2}\varphi_h\big\|_{n-1/2}$$

with a constant C independent of h, τ and n^*.

Proof. First note that

$$\big\|\partial_\tau(\tfrac{3}{2}u_h^k - \tfrac{1}{2}u_h^{k-1})\big\|_{L^\infty} \le \tfrac{3}{2}\left\|\partial_\tau u_h^k\right\|_{L^\infty} + \tfrac{1}{2}\left\|\partial_\tau u_h^{k-1}\right\|_{L^\infty},$$

as well as

$$\big\|\tfrac{3}{2}u_h^k - \tfrac{1}{2}u_h^{k-1}\big\|_{L^\infty} = \big\|u_h^k + \tfrac{\tau}{2}\partial_\tau u_h^k\big\|_{L^\infty}.$$

This allows us to proceed as in Lemma 4.4 and to bound $\partial_\tau^j\Lambda_h(\bar{y}^{n+1/2})$, $j = 0, 1$, and the inverse $\Lambda_h^{-1}(\bar{y}^{n+1/2})$. □

We are then able to conclude the error bound for the midpoint rule.

Proof of Proposition 4.5. Using the error equation (4.12), we employ Lemmas 4.6 and 4.7 and (4.10) to obtain

$$\begin{aligned}\big\|\mathbf{A}_h^{n+1/2}e_h^{n+1}\big\|_{n+1/2} &\leq (1+C\tau)\big\|\mathbf{A}_h^{n-1/2}e_h^{n}\big\|_{n-1/2}\\ &\quad + C\tau\big(\big\|\mathrm{A}_h e_h^n\big\|_{X_h} + \big\|\mathrm{A}_h e_h^{n-1}\big\|_{X_h}\big) + C\tau\big(\tau^2+h^k\big).\end{aligned}$$

With the bound on $\mathrm{A}_h e_h^0$ from Proposition 3.5 and using the fact the first step is given by the Euler method, we obtain with Proposition 4.2

$$\big\|\mathrm{A}_h e_h^1\big\|_{X_h} \leq C\tau\big(\tau+h^k\big),$$

which yields by a Gronwall lemma the assertion. □

4.3. **Exponential Euler.** For the exponential method, we apply a similar approach and derive the necessary bound in the stronger energy norm. However, there is no direct relation to the discrete derivatives of the error. In this case, we have to prove an additional error estimate.

Proposition 4.8. *Under the assumptions of Theorem 2.8, for* $0 \leq n \leq n^*+1$ *there hold the bounds*

$$\big\|\mathrm{A}_h e_h^n\big\|_{X_h} \leq C\big(\tau+h^k\big),$$

and for $1 \leq n \leq n^*+1$

$$\big\|\partial_\tau e_h^n\big\|_{X_h} \leq C\big(\tau+h^k\big),$$

where C *is independent of* h, τ *and* n^*.

Once this is established, the last main result directly follows.

Proof of Theorem 2.8. In order to conclude the convergence rates, we only employ the first estimate in Proposition 4.8. To show $n^* \geq N-1$, again the first estimate allows us to guarantee the first three bounds in (4.1). The bound on $\mathcal{L}_h\partial_\tau u_h^k$ follows from the second estimate in Proposition 4.8 combined with Lemma 4.1 and the CFL-type condition (2.18). □

The rest of this section is devoted to the proof of Proposition 4.8. We introduce the auxiliary approximation $\tilde{y}^n(t^n+s)$ for $s \in [0,\tau]$ as the solution of

$$\Lambda_h(y_h^n)\partial_t\tilde{y}^n(t^n+s) = \mathrm{A}_h\tilde{y}^n(t^n+s) + G_h(t^n,y_h^n), \quad \tilde{y}^n(t^n) = y_h^n \tag{4.13}$$

and thus satisfies $\tilde{y}^n(t^n+\tau) = y_h^{n+1}$. In order to derive the error equation, we insert the projected exact solution $J_h y$ in (4.13)

$$\begin{aligned}\Lambda_h(y_h^n)J_h\partial_t y(t^n+s) &= \mathrm{A}_h J_h y(t^n+s) + G_h(t^n, J_h y(t^n))\\ &\quad + \big(\Lambda_h(y_h^n) - \Lambda_h(J_h y(t^n))\big)J_h\partial_t y(t^n+s) + \delta_{\mathrm{ExEu}}^{n+1}(t^n+s)\end{aligned}$$

with defect $\delta_{\mathrm{ExEu}}^{n+1} = \delta_{h,\mathrm{ExEu}}^{n+1} + \delta_{\tau,\mathrm{ExEu}}^{n+1}$ given by

$$\begin{aligned}\delta_{h,\mathrm{ExEu}}^{n+1}(t^n+s) &= \big(J_h\mathrm{A} - \mathrm{A}_h J_h\big)y(t^n+s) + J_h G(t^n,y(t^n)) - G_h(t^n, J_h y(t^n))\\ &\quad + \big(\Lambda_h(J_h y(t^n))J_h - J_h\Lambda(y(t^n))\big)\partial_t y(t^n+s),\\ \delta_{\tau,\mathrm{ExEu}}^{n+1}(t^n+s) &= J_h\Lambda(y(t^n))\partial_t y(t^n+s) - < J_h\Lambda(y(t^n+s)\partial_t y(t^n+s)\\ &\quad + J_h G(t^n+s, y(t^n+s)) - J_h G(t^n, y(t^n)).\end{aligned}$$

Similarly, we define the auxiliary error by

$$\tilde{e}_h^n(t^n+s) \coloneqq J_h y(t^n+s) - \tilde{y}^n(t^n+s), \quad \tilde{e}_h^n(t^n) = e_h^n, \quad \tilde{e}_h^n(t^n+\tau) = e_h^{n+1}.$$

This yields the discrete error equation for $s \in [0,\tau]$

$$\Lambda_h(y_h^n)\partial_t \tilde{e}_h^n(t^n+s) = \mathrm{A}_h \tilde{e}_h^n(t^n+s) + \Gamma_h^n(t^n+s) + \delta_{\mathrm{ExEu}}^{n+1}(t^n+s) \tag{4.14}$$

with stability term

$$\Gamma_h^n(t^n+s) \coloneqq \big(G_h(t^n, J_h y(t^n)) - G_h(t^n, y_h)\big) + \big(\Lambda_h(y_h^n) - \Lambda_h(J_h y(t^n))\big) J_h \partial_t y((t^n+s)).$$

Using the variation-of-constants formula with the state-dependent operator defined in (4.6), we obtain from (4.14)

$$\tilde{e}_h^n(t^n+s) = \mathrm{e}^{s\mathbf{A}_h^n} e_h^n + \int_0^s \mathrm{e}^{(s-\sigma)\mathbf{A}_h^n} \Lambda_h^{-1}(y_h^n)\big(\Gamma_h^n(t^n+\sigma) + \delta_{\mathrm{ExEu}}^{n+1}(t^n+\sigma)\big)\, d\sigma.$$

To obtain the error bounds stated in Proposition 4.8, we need the following two estimates which follow along the lines of Lemmas 3.7 and 4.3: For $0 \le n \le n^*$ it holds

$$\sup_{s\in[0,\tau]} \|\mathrm{A}_h \Gamma_h^n(t^n+s)\|_{X_h} \le C \|\mathrm{A}_h e_h^n\|_{X_h}, \tag{4.15a}$$

$$\sup_{s\in[0,\tau]} \big\|\mathrm{A}_h \delta_{\mathrm{ExEu}}^{n+1}(t^n+s)\big\|_{X_h} \le C\big(\tau + h^k\big), \tag{4.15b}$$

with a constant C independent of h, τ and n^*. This allows us to conclude the bounds in the two stronger norms.

Proof of Proposition 4.8. We proceed as in the proof of Proposition 4.2 in order to obtain the bound on $\mathrm{A}_h \tilde{e}_h^n$ in the form

$$\sup_{s\in[0,\tau]} \big\|\mathrm{A}_h \tilde{e}_h^n(t^n+s)\big\|_{X_h} \le C\big(\tau + h^k\big), \tag{4.16}$$

which implies the first statement in the proposition. For the discrete derivative of the error we employ (4.7), (4.14), and (4.15) to conclude

$$\begin{aligned}
\big\|\partial_\tau e_h^n\big\|_{X_h} &\le \sup_{s\in[0,\tau]} \big\|\partial_t \tilde{e}_h^n(t^n+s)\big\|_{X_h} \\
&\le C \sup_{s\in[0,\tau]} \big\|\mathrm{A}_h \tilde{e}_h^n(t^n+s)\big\|_{X_h} + C\|\mathrm{A}_h e_h^n\|_{X_h} + C\big(\tau + h^k\big) \\
&\le C\big(\tau + h^k\big),
\end{aligned}$$

where we used (4.16) in the last step. □

4.4. **Exponential midpoint rule.** For the exponential midpoint rule, we combine the approaches presented for the semi-implicit midpoint rule and the exponential Euler method. In particular, we have to prove error bounds in the stronger norm as well as for the discrete derivative of the error.

Proposition 4.9. *Under the assumptions of Theorem 2.9, for $0 \le n \le n^*+1$ there hold the bounds*

$$\big\|\mathrm{A}_h e_h^n\big\|_{X_h} \le C\big(\tau^2 + h^k\big),$$

and for $1 \le n \le n^*+1$

$$\big\|\partial_\tau e_h^n\big\|_{X_h} \le C\big(\tau^2 + h^k\big),$$

where C is independent of h, τ and n^*.

Once this is established, the last main result directly follows.

Proof of Theorem 2.9. We only combine the argument presented in the proofs of Theorems 2.7 and 2.8 to conclude the assertion. $\square$

The rest of this section is devoted to the proof of Proposition 4.9. We introduce the auxiliary approximation $\tilde{y}^n(t^n+s)$ for $s \in [0,\tau]$ as the solution of

$$\Lambda_h(\bar{y}_h^{n+1/2})\partial_t \tilde{y}^n(t^n+s) = \mathrm{A}_h \tilde{y}^n(t^n+s) + G_h(t^{n+1/2}, \bar{y}_h^{n+1/2}) \tag{4.17}$$

with $\tilde{y}^n(t^n) = y_h^n$, and thus satisfies $\tilde{y}^n(t^n+\tau) = y_h^{n+1}$. In order to derive the error equation, we insert the projected exact solution $J_h y$ in (4.17) and conclude

$$\begin{aligned}\Lambda_h(\bar{y}_h^{n+1/2}) J_h \partial_t y(t^n+s) &= \mathrm{A}_h J_h y(t^n+s) + G_h(t^{n+1/2}, J_h \bar{y}^{n+1/2}) \\ &\quad + \big(\Lambda_h(\bar{y}_h^{n+1/2}) - \Lambda_h(J_h \bar{y}^{n+1/2})\big) J_h \partial_t y(t^n+s) + \delta_{\mathrm{ExM}}^{n+1}(t^n+s)\end{aligned}$$

with defect $\delta_{\mathrm{ExM}}^{n+1} = \delta_{h,\mathrm{ExM}}^{n+1} + \delta_{\tau,\mathrm{ExM},1}^{n+1} + \delta_{\tau,\mathrm{ExM},2}^{n+1}$ given by

$$\begin{aligned}\delta_{h,\mathrm{ExM}}^{n+1}(t^n+s) &= \big(J_h \mathrm{A} - \mathrm{A}_h J_h\big) y(t^n+s) + J_h G(t^{n+1/2}, \bar{y}^{n+1/2}) - G_h(t^{n+1/2}, J_h \bar{y}^{n+1/2}) \\ &\quad + \big(\Lambda_h(J_h \bar{y}^{n+1/2}) J_h - J_h \Lambda(\bar{y}^{n+1/2})\big) \partial_t y(t^n+s), \\ \delta_{\tau,\mathrm{ExM},1}^{n+1}(t^n+s) &= J_h \Lambda(\bar{y}^{n+1/2}) \partial_t y(t^n+s) - J_h \Lambda(y(t^{n+1/2})) \partial_t y(t^n+s) \\ &\quad + J_h G(t^{n+1/2}, y(t^{n+1/2})) - J_h G(t^{n+1/2}, J_h \bar{y}^{n+1/2}), \\ \delta_{\tau,\mathrm{ExM},2}^{n+1}(t^n+s) &= J_h \Lambda(y(t^{n+1/2})) \partial_t y(t^n+s) - J_h \Lambda(y(t^n+s)) \partial_t y(t^n+s) \\ &\quad + J_h G(t^n+s, y(t^n+s)) - J_h G(t^{n+1/2}, y(t^{n+1/2})).\end{aligned}$$

Deriving the error equation and using the variation-of-constants formula, we obtain with the state-dependent operator $\mathbf{A}_h^{n+1/2} = \Lambda_h^{-1}(\bar{y}_h^{n+1/2}) \mathrm{A}_h$

$$e_h^{n+1} = \mathrm{e}^{\tau \mathbf{A}_h^{n+1/2}} e_h^n + \int_0^\tau \mathrm{e}^{(\tau-\sigma)\mathbf{A}_h^{n+1/2}} \Lambda_h^{-1}(\bar{y}_h^{n+1/2}) \big(\Gamma_h^n(t^n+\sigma) + \delta_{\mathrm{ExM}}^{n+1}(t^n+\sigma)\big) \, d\sigma$$

with stability term

$$\begin{aligned}\Gamma_h^n(t^n+s) \coloneqq\ & G_h(t^{n+1/2}, J_h \bar{y}^{n+1/2}) - G_h(t^{n+1/2}, \bar{y}_h^{n+1/2}) \\ & + \big(\Lambda_h(\bar{y}_h^{n+1/2}) - \Lambda_h(J_h \bar{y}^{n+1/2})\big) J_h \partial_t y(t^n+s).\end{aligned}$$

Unlike for the exponential Euler method, one has to pay more attention to the derivation of the error bound on the discrete derivatives. In order to show the error bound, we do not only apply $\mathbf{A}_h^{n+1/2}$ to the error equation, but also apply the discrete derivative ∂_τ. In a straightforward manner, one can derive the following auxiliary result.

Lemma 4.10. *Under the assumptions of Theorem 2.9, it holds for $0 \le \sigma \le \tau$*

$$\big\| \partial_\tau \big(\mathrm{e}^{\sigma \mathbf{A}_h^{n+1/2}} \varphi_h^n\big) - \mathrm{e}^{\sigma \mathbf{A}_h^{n+1/2}} \partial_\tau \varphi_h^n \big\|_{X_h} \le C\tau \, \|\mathrm{A}_h \varphi_h^n\|_{X_h}$$

with a constant C independent of h, τ and n^*.

This leads to several error terms which appeared above similarly. Along the lines of Lemmas 3.7, 4.3, and 4.6 we immediately conclude the following bounds: For

$0 \le n \le n^*$ the stability terms are bounded by

$$\sup_{s\in[0,\tau]} \|\mathrm{A}_h \Gamma_h^n(t^n+s)\|_{X_h} \le C\big(\|\mathrm{A}_h e_h^n\|_{X_h} + \big\|\mathrm{A}_h e_h^{n-1}\big\|_{X_h}\big), \tag{4.18a}$$

$$\sup_{s\in[0,\tau]} \|\partial_\tau \Gamma_h^n(t^n+s)\|_{X_h} \le C \sum_{j\in\{n-1,n\}} \big\|\partial_\tau e_h^j\big\|_{X_h} + \big\|e_h^j\big\|_{X_h}, \tag{4.18b}$$

and further the defects satisfy

$$\sup_{s\in[0,\tau]} \big\|\mathrm{A}_h \delta_{h,\mathrm{ExM}}^{n+1}(t^n+s)\big\|_{X_h} + \sup_{s\in[0,\tau]} \big\|\partial_\tau \delta_{h,\mathrm{ExM}}^{n+1}(t^n+s)\big\|_{X_h} \le Ch^k, \tag{4.19a}$$

$$\sup_{s\in[0,\tau]} \big\|\mathrm{A}_h \delta_{\tau,\mathrm{ExM},1}^{n+1}(t^n+s)\big\|_{X_h} + \sup_{s\in[0,\tau]} \big\|\partial_\tau \delta_{\tau,\mathrm{ExM},1}^{n+1}(t^n+s)\big\|_{X_h} \le C\tau^2, \tag{4.19b}$$

$$\sup_{s\in[0,\tau]} \big\|\mathrm{A}_h \delta_{\tau,\mathrm{ExM},2}^{n+1}(t^n+s)\big\|_{X_h} \le C\tau, \tag{4.19c}$$

with a constant C independent of h, τ and n^*. The main difficulty is to extract the additional order of convergence in the defect $\delta_{\tau,\mathrm{ExM},2}^n$. We show in Section 5 the following lemma.

Lemma 4.11. *Under the assumptions of Theorem 2.9 it holds*

$$\Big\|\mathbf{A}_h^{n+1/2} \int_0^\tau \mathrm{e}^{(\tau-\sigma)\mathbf{A}_h^{n+1/2}} \Lambda_h^{-1}(\bar{y}_h^{n+1/2}) \delta_{\tau,ExM,2}^{n+1}(t^n+\sigma)\big)\, d\sigma\Big\|_{X_h} \le C\tau^3,$$
$$\Big\|\int_0^\tau \mathrm{e}^{(\tau-\sigma)\mathbf{A}_h^{n+1/2}} \Lambda_h^{-1}(\bar{y}_h^{n+1/2}) \partial_\tau \delta_{\tau,ExM,2}^{n+1}(t^n+\sigma)\big)\, d\sigma\Big\|_{X_h} \le C\tau^3,$$

with a constant C independent of h, τ and n^*.

From this, we conclude the error bounds in Proposition 4.9.

Proof of Proposition 4.9. Following the lines of the preceding proofs of Propositions 4.5 and 4.8, one establishes the error bound on $\|\mathrm{A}_h e_h^n\|_{X_h}$. Applying the discrete derivative to the error equation and employing the bounds in Lemma 4.10 combined with (4.18) and (4.19) yields

$$\partial_\tau e_h^{n+1} = \mathrm{e}^{\tau \mathbf{A}_h^{n+1/2}} \partial_\tau e_h^n + \int_0^\tau \mathrm{e}^{(\tau-\sigma)\mathbf{A}_h^{n+1/2}} \Lambda_h^{-1}(\bar{y}_h^{n+1/2}) \partial_\tau \delta_{\mathrm{ExM}}^{n+1}(t^n+\sigma)\, d\sigma + \Delta_h^n,$$

where the remainder term Δ_h^n satisfies

$$\|\Delta_h^n\|_{X_h} \le C\tau\big(\|\partial_\tau e_h^n\|_{X_h} + \big\|\partial_\tau e_h^{n-1}\big\|_{X_h}\big) + C\tau\big(\tau^2 + h^k\big).$$

Finally, the application of Lemma 4.11 yields the desired estimate on $\|\partial_\tau e_h^n\|_{X_h}$ and closes the proof. □

5. Estimates for stability terms and defects

This section is devoted to the proofs of the postponed stability and consistency estimate from Sections 3 and 4.

5.1. **Stability.** In the following, we give a detailed proof for the stability term given in (3.10). We emphasize that the corresponding bounds used in Section 4, are derived fully analogously, and we thus refrain from giving the details here.

Proof of Lemma 3.7. We consider the two contributions of Γ_h in (3.10) separately.

(a) We first note by (2.6) and (2.7) that

$$\mathrm{A}_h\big(G_h(t,J_h y)-G_h(t,y_h)\big)=\begin{pmatrix} I_h g(t,\mathcal{L}_h\mathcal{L}_h^{V*}u,\mathcal{L}_h\mathcal{L}_h^{V*}\partial_t u)-I_h g(t,\mathcal{L}_h u_h,\mathcal{L}_h v_h)\\ 0\end{pmatrix}.$$

Without loss of generality, we show the assertion only for $g(t,u,\partial_t u)=g(\partial_t u)$, and obtain

$$\begin{aligned}
&\big\|\mathrm{A}_h\big(G_h(t,J_h y)-G_h(t,y_h)\big)\big\|_{X_h}\\
&=\big\|I_h g(\mathcal{L}_h\mathcal{L}_h^{V*}\partial_t u)-I_h g(\mathcal{L}_h v_h)\big\|_{V_h}\\
&=\Big\|I_h\Big(\int_0^1 g'(\sigma\mathcal{L}_h\mathcal{L}_h^{V*}\partial_t u+(1-\sigma)\mathcal{L}_h v_h)\,d\sigma\,(\mathcal{L}_h\mathcal{L}_h^{V*}\partial_t u-\mathcal{L}_h v_h)\Big)\Big\|_{V_h}\\
&\lesssim\Big\|\int_0^1 g'(\sigma\mathcal{L}_h\mathcal{L}_h^{V*}\partial_t u+(1-\sigma)\mathcal{L}_h v_h)\,d\sigma\Big\|_{W^{1,N+\delta}}\big\|\mathcal{L}_h\mathcal{L}_h^{V*}\partial_t u-\mathcal{L}_h v_h\big\|_{V_h},
\end{aligned}$$

where we used Lemma 2.10 for the last estimate. The latter term is estimated with (3.4) by $\mathrm{A}_h e_h$ in the X_h-norm. For the integral part, we use the stability of $\mathcal{L}_h\mathcal{L}_h^{V*}$ in (2.24) with $\ell=0$ to bound it by a constant depending on the $W^{1,N+\delta}$-norms of $\partial_t u$ and v_h. Hence, the bounds in (3.3) yield the stability for G_h.

(b) Next, we consider by (2.6) and (2.7)

$$\mathrm{A}_h\big(\Lambda_h(J_h y(t))-\Lambda_h(y_h(t))J_h\partial_t y(t)=\begin{pmatrix}\pi_h\Big(I_h\big(\lambda(\mathcal{L}_h\mathcal{L}_h^{V*}u)-\lambda(\mathcal{L}_h u_h)\big)\mathcal{L}_h^{V*}\partial_t^2 u\Big)\\ 0\end{pmatrix},$$

and estimate with the stability of the L^2-projection in (2.26) and of the interpolation (2.22), the algebra property of $W^{1,N+\delta}$, and the stability of $\mathcal{L}_h^{V*}$ in (2.24)

$$\begin{aligned}
&\big\|\mathrm{A}_h\big(\Lambda_h(J_h y(t))-\Lambda_h(y_h(t))J_h\partial_t y(t)\big\|_{X_h}\\
&\le C\,\big\|\big(\lambda(\mathcal{L}_h\mathcal{L}_h^{V*}u)-\lambda(\mathcal{L}_h u_h)\big\|_{W^{1,N+\delta}}\,\big\|\mathcal{L}_h^{V*}\partial_t^2 u\big\|_{W^{1,N+\delta}}\\
&\le C\,\big\|\mathcal{L}_h^{V*}u-u_h\big\|_{W^{1,N+\delta}}\,\big\|\partial_t^2 u\big\|_{W^{1,N+\delta}},
\end{aligned}$$

with a constant depending on the $W^{1,N+\delta}$-norms of u and u_h. Using Lemma 3.3 and (3.4), the first term is bounded by $\mathrm{A}_h e_h$ in the X_h-norm. □

5.2. **Defects.** We first estimate the spatial defect (3.8) which will reappear in a modified form in the defects of the full discretization.

Proof of Lemma 3.8. (a) We compute with (2.27) and (2.28)

$$\mathrm{A}_h\big(J_h\mathrm{A}-\mathrm{A}_h J_h\big)y(t)=\mathrm{A}_h\big(J_h-\boldsymbol{\mathcal{L}}_h^*\big)\mathrm{A}y(t)=\begin{pmatrix}(\mathcal{L}_h^{V*}-\mathcal{L}_h^{H*})\Delta u\\ 0\end{pmatrix}$$

and, inserting the interpolation, estimate with (2.22), (2.23), (2.24), and (2.25),

$$\big\|\mathrm{A}_h\big(J_h\mathrm{A}-\mathrm{A}_h J_h\big)y(t)\big\|_{X_h}\le Ch^k\,\|\Delta u\|_{H^{k+1}}\,.$$

(b) As above, we only consider the case $g(t,u,\partial_t u) = g(\partial_t u)$ and obtain with (2.6) and (2.7)

$$\mathrm{A}_h\big(J_h G(t,y) - G_h(t,J_h y)\big) = \begin{pmatrix} \mathcal{L}_h^{V*} g(\partial_t u) - I_h g(\mathcal{L}_h \mathcal{L}_h^{V*}\partial_t u) \\ 0 \end{pmatrix}.$$

From this, we conclude with (2.22) and (2.24)

$$\begin{aligned}
&\big\|\mathrm{A}_h\big(J_h G(t,y) - G_h(t,J_h y)\big)\big\|_{X_h} \\
&\le \big\|(\mathcal{L}_h^{V*} - I_h) g(\partial_t u)\big\|_{V_h} + \big\|I_h g(\partial_t u) - I_h g(\mathcal{L}_h \mathcal{L}_h^{V*}\partial_t u)\big\|_{V_h} \\
&\le C h^k \left\|g(\partial_t u)\right\|_{H^{k+1}} + C \big\|(\mathrm{Id} - \mathcal{L}_h \mathcal{L}_h^{V*})\partial_t u\big\|_{W^{1,\infty}} \\
&\le C(\left\|g(\partial_t u)\right\|_{H^{k+1}} + \left\|\partial_t u\right\|_{W^{k+1,\infty}})\, h^k,
\end{aligned}$$

which gives the desired convergence rate.

(c) We compute with (2.6) and (2.7)

$$\mathrm{A}_h\big(J_h\Lambda(y) - \Lambda_h(J_h y) J_h\big)\partial_t y = \begin{pmatrix} \mathcal{L}_h^{V*}(\lambda(u)\partial_t^2 u) - \pi_h\big(I_h\lambda(\mathcal{L}_h\mathcal{L}_h^{V*}u)\mathcal{L}_h\mathcal{L}_h^{V*}\partial_t^2 u\big) \\ 0 \end{pmatrix},$$

such that again (2.22), (2.23) and (2.24) yield the estimate

$$\begin{aligned}
&\big\|\mathrm{A}_h\big(J_h\Lambda(y) - \Lambda_h(J_h y)J_h\big)\partial_t y\big\|_{X_h} \\
&\le \big\|(\mathcal{L}_h^{V*} - \mathcal{L}_h^{H*})\lambda(u)\partial_t^2 u\big\|_{V_h} + \big\|(\mathcal{L}_h^{H*} - \pi_h\mathcal{L}_h^{-1})\lambda(u)\partial_t^2 u\big\|_{V_h} \\
&\qquad + \big\|\lambda(u)\partial_t^2 u - \mathcal{L}_h I_h \lambda(\mathcal{L}_h\mathcal{L}_h^{V*}u)\mathcal{L}_h\mathcal{L}_h^{V*}\partial_t^2 u\big\|_{V_h} \\
&\le C(\left\|\lambda(u)\right\|_{W^{k+1,\infty}}, \left\|\partial_t^2 u\right\|_{W^{k+1,\infty}})\, h^k + h^{-1}\big\|(\mathcal{L}_h^{H*} - \pi_h\mathcal{L}_h^{-1})\lambda(u)\partial_t^2 u\big\|_{H_h}.
\end{aligned}$$

The last term is estimated using [16, Lem. 8.24] to obtain

$$\begin{aligned}
\big\|(\mathcal{L}_h^{H*} - \pi_h\mathcal{L}_h^{-1})\lambda(u)\partial_t^2 u\big\|_{H_h} &\lesssim h^k \big\|\mathcal{L}_h^{-1}\lambda(u)\partial_t^2 u\big\|_{H_h} \\
&\lesssim h^k \big\|(\mathcal{L}_h^{-1} - I_h)\lambda(u)\partial_t^2 u\big\|_{H_h} + h^k \big\|I_h\lambda(u)\partial_t^2 u\big\|_{H_h},
\end{aligned}$$

and (2.22) together with (B.2) gives the desired bound. □

For the fully discrete defects, we rely on further Lipschitz bounds of the nonlinearities Λ and G which we collect in the next lemma. Since, we work in a first-order framework we denote in the following for any function $x \in X$, the projection onto the first and second component by x_1 or x_2, respectively.

Lemma 5.1. *Let $x, y, z \in X$, and let Assumption 2.1 hold.*

(a) If $x_1, y_1, z_2 \in W^{1,\infty}(\Omega)$, then

$$\big\|\mathrm{A}\big(\Lambda(x) - \Lambda(y)\big)z\big\|_X \le C \left\|x_1 - y_1\right\|_{H^1(\Omega)},$$

where the constant depends on the $W^{1,\infty}$-norms of x_1, y_1, z_2.

(b) If $x_1, x_2, y_1, y_2 \in W^{1,\infty}(\Omega)$, then

$$\big\|\mathrm{A}\big(G(t,x) - G(s,y)\big)\big\|_X \le C\big(|t-s| + \left\|x_1 - y_1\right\|_{H^1(\Omega)} + \left\|x_2 - y_2\right\|_{H^1(\Omega)}\big)$$

where the constant depends on the $W^{1,\infty}$-norms of x_1, x_2, y_1, y_2.

Proof. We expand the difference in part (a) as

$$\begin{aligned}\left\|\mathrm{A}\big(\Lambda(x)-\Lambda(y)\big)z\right\|_X &= \left\|\big(\lambda(x_1)-\lambda(y_1)\big)z_2\right\|_{H^1}\\ &= \left\|\int_0^1 \lambda'(\sigma x_1+(1-\sigma y_1))(x_1-y_1)z_2\right\|_{H^1}\\ &\lesssim \sup_{\sigma\in[0,1]}\left\|\lambda'(\sigma x_1+(1-\sigma y_1))\right\|_{W^{1,\infty}}\|x_1-y_1\|_{H^1}\|z_2\|_{W^{1,\infty}},\end{aligned}$$

and the assumptions in the lemma yield the bound. The very same computation yields the second estimate (b). □

Thus, the defect of the Euler method can be bounded in a straightforward way.

Proof of Lemma 4.3. We recall the splitting $\delta_{\mathrm{Eu}}^{n+1}=\delta_{h,\mathrm{Eu}}^{n+1}+\delta_{\tau,\mathrm{Eu}}^{n+1}$ of the defect in (4.3), and note that by proof of Lemma 3.8 it holds

$$\left\|\mathrm{A}_h\delta_{h,\mathrm{Eu}}^{n+1}\right\|_{X_h}\le Ch^k.$$

We treat the two parts in (4.3b) separately. The second term involving G is bounded using (2.28) and Lemma 5.1. Further, we expand

$$\begin{aligned}&\mathrm{A}_hJ_h\Lambda(y(t^n))\partial_\tau y(t^{n+1})-\mathrm{A}_hJ_h\Lambda(y(t^{n+1}))\partial_t y(t^{n+1})\\ &=\boldsymbol{\mathcal{L}}_{\boldsymbol{h}}^*\mathrm{A}\big(\Lambda(y(t^n))-\Lambda(y(t^{n+1}))\big)\partial_t y(t^{n+1})+\boldsymbol{\mathcal{L}}_{\boldsymbol{h}}^*\mathrm{A}\Lambda(y(t^n))\big(\partial_\tau y(t^{n+1})-\partial_t y(t^{n+1})\big),\end{aligned}$$

such that Lemma 5.1 is employed on the first part. For the second part, note that by the fundamental theorem of calculus we obtain in any norm

$$\left\|\partial_\tau z(t^{n+1})-\partial_t z(t^{n+1})\right\|\le\frac{\tau}{2}\sup_{s\in[0,\tau]}\left\|\partial_t^2 z(t^n+s)\right\|,$$

and the claim follows. □

Similarly, we bound the defect of the midpoint rule in (4.11), and as above we do not have to treat the spatial part $\delta_{h,\mathrm{M}}^{n+1}$.

Proof of Lemma 4.6. We first note that from Taylor expansions and the Peano kernel theorem, we conclude in any norm the bounds

$$\begin{aligned}\left\|\partial_\tau z(t^{n+1})-\partial_t z(t^{n+1/2})\right\|&\le\frac{\tau^2}{24}\sup_{s\in[0,\tau]}\left\|\partial_t^3 z(t^n+s)\right\|,\\ \left\|z(t^{n+1/2})-\tfrac12\big(z(t^{n+1})+z(t^n)\big)\right\|&\le\frac{\tau^2}{4}\sup_{s\in[0,\tau]}\left\|\partial_t^2 z(t^n+s)\right\|,\\ \left\|z(t^{n+1/2})-\big(\tfrac32 z(t^n)-\tfrac12 z(t^{n-1})\big)\right\|&\le\frac{3\tau^2}{8}\sup_{s\in[0,\tau]}\left\|\partial_t^2 z(t^n+s)\right\|.\end{aligned}$$

Combining this with Lemma 5.1 and the proof of Lemma 4.3 yields the desired bounds on the defect. □

We finally treat the principle defect of the exponential midpoint rule. We pursue the strategy adapted from [12, Prop. 5.3]. Before we give the proof, we need two auxiliary results. The first one allows us to compare function evaluations of finite element objects with their interpolation. The proof is given in Appendix B.

Lemma 5.2. *Let $L \in \mathbb{N}$ and assume that $f\colon \mathbb{R}^L \to \mathbb{R}$ is sufficiently often differentiable, and that $\varphi_{i,h} \in V_h$ satisfies*

$$\|\varphi_{i,h}\|_{L^\infty} + \|\varphi_{i,h}\|_{W^{1,4}} \le C$$

for $i = 1, \dots. L$. Then

$$\|f(\varphi_{1,h}, \dots, \varphi_{L,h}) - I_h f(\mathcal{L}_h\varphi_{1,h}, \dots, \mathcal{L}_h\varphi_{L,h})\|_{L^2(\Omega_h)} \lesssim h^2,$$
$$\|f(\varphi_{1,h}, \dots, \varphi_{L,h}) - I_h f(\mathcal{L}_h\varphi_{1,h}, \dots, \mathcal{L}_h\varphi_{L,h})\|_{H^1(\Omega_h)} \lesssim h,$$

with a constant independent of h.

On the continuous level, the chain rule allows us to bound terms of the form $\Delta(\mu(u)w)$ by the norms of Δu and Δw. Even though, this is not straightforward in the discrete case, one can establish a very similar result. We recall the Ritz projection R_h defined in (2.5), and note that the proof is given in Appendix D.

Lemma 5.3. *Let $u_h, w_h \in V_h$ and assume that $\|\Delta_h u_h\|_{L^2} + \|\Delta_h w_h\|_{L^2} \le C$. Further, let $\mu\colon \mathbb{R} \to \mathbb{R}$ be twice continuously differentiable. Then*

$$\|\Delta_h R_h(\mu(u_h)w_h)\|_{L^2(\Omega_h)} \le C$$

with a constant C independent of h.

With these preparations, we can provide the error bounds on the defects of the exponential midpoint rule.

Proof of Lemma 4.11. Let us denote

$$d_n^1(t) := J_h\big(\Lambda(y(t^{n+1/2})) - \Lambda(y(t))\big)\partial_t y(t),$$
$$d_n^2(t) := J_h\big(G(t, y(t)) - G(t^{n+1/2}, y(t^{n+1/2}))\big),$$

then it remains to show for $i = 1, 2$

$$\Big\|\mathbf{A}_h^{n+1/2}\int_0^\tau \mathrm{e}^{(\tau-\sigma)\mathbf{A}_h^{n+1/2}}\Lambda_h^{-1}(\bar{y}_h^{n+1/2})d_n^i(t^n+\sigma)\big)\,d\sigma\Big\|_{X_h} \le C\tau^3,$$
$$\Big\|\int_0^\tau \mathrm{e}^{(\tau-\sigma)\mathbf{A}_h^{n+1/2}}\Lambda_h^{-1}(\bar{y}_h^{n+1/2})\partial_\tau d_n^i(t^n+\sigma)\big)\,d\sigma\Big\|_{X_h} \le C\tau^3.$$

Following the lines of the proof of [12, Prop. 5.3], we observe that for the first bound on d_n^1, we have to establish the bounds

$$\big\|\mathrm{A}_h\mathbf{A}_h^{n+1/2}\Lambda_h^{-1}(\bar{y}_h^{n+1/2})J_h\big(\partial_t\Lambda(y(t))\big)\partial_t y(t)\big\|_{X_h} \le C, \tag{5.1a}$$
$$\big\|\mathrm{A}_h\Lambda_h^{-1}(\bar{y}_h^{n+1/2})J_h\big(\Lambda(y(t^{n+1/2})) - \Lambda(y(t))\big)\partial_t^2 y(t)\big\|_{X_h} \le C\tau, \tag{5.1b}$$
$$\big\|\mathrm{A}_h\Lambda_h^{-1}(\bar{y}_h^{n+1/2})J_h\Big(\big(\partial_t^2\Lambda(y(t))\big)\partial_t y(t) + \big(\partial_t\Lambda(y(t))\big)\partial_t^2 y(t)\Big)\big\|_{X_h} \le C, \tag{5.1c}$$

as well as for d_n^2 the bounds

$$\big\|\mathrm{A}_h\mathbf{A}_h^{n+1/2}\Lambda_h^{-1}(\bar{y}_h^{n+1/2})J_h\partial_t G(t, y(t))\big\|_{X_h} \le C, \tag{5.2a}$$
$$\big\|\mathrm{A}_h\Lambda_h^{-1}(\bar{y}_h^{n+1/2})J_h\partial_t^2 G(t, y(t))\big\|_{X_h} \le C. \tag{5.2b}$$

For the discrete derivative, we have similar terms which have one A_h less, and instead ∂_τ applied to objects following $\Lambda_h^{-1}(\bar{y}_h^{n+1/2})$. In the following, we only discuss the bounds in (5.1a) and (5.2a) since the remaining ones are more standard.

Denoting $w_h := \mathcal{L}_h^{V*}\big(\lambda'(u)\,\partial_t u\,\partial_t^2 u\big)$, it is sufficient to show for the bound in (5.1a) the estimate

$$\big\|\Delta_h Q_h\big((I_h\lambda(\mathcal{L}_h\bar{u}_h^{n+1/2})^{-1})w_h\big)\big\|_{H_h} \le C,$$

since the first component vanishes. In the following, we will often use that by (2.28) and the assumptions of Theorem 2.9 it holds

$$\|\Delta_h w_h\|_{L^2(\Omega_h)} = \big\|\mathcal{L}_h^{H*}\Delta\big(\lambda'(u)\,\partial_t u\,\partial_t^2 u\big)\big\|_{L^2(\Omega_h)} \le C,$$

and to keep the notation simple, we will write u_h instead of $\bar{u}_h^{n+1/2}$. We split the term using the inverse estimate in the form (A.1) in

$$\begin{aligned}\big\|\Delta_h Q_h\big((I_h\lambda(\mathcal{L}_h u_h))^{-1}w_h\big)\big\|_{H_h} \lesssim{}& \big\|\Delta_h R_h\big(\lambda(u_h)^{-1}w_h\big)\big\|_{H_h} \\ &+ h^{-1}\big\|Q_h\big((I_h\lambda(\mathcal{L}_h u_h))^{-1}w_h\big) - R_h\big(\lambda(u_h)^{-1}w_h\big)\big\|_{V_h}.\end{aligned}$$

The first term is bounded by Lemma 5.3 using

$$\big\|\Delta_h\bar{u}_h^{n+1/2}\big\|_{L^2(\Omega_h)} \lesssim \|\Delta_h u_h^n\|_{L^2(\Omega_h)} + \big\|\Delta_h u_h^{n-1}\big\|_{L^2(\Omega_h)}.$$

To split the second term further, we add and subtract in the second term $Q_h\big(\lambda(u_h)^{-1}w_h\big)$, and obtain for the first term with the L^2-stability of Q_h in (2.26) and the inverse estimate (2.25)

$$\begin{aligned}&h^{-1}\big\|Q_h\big(((I_h\lambda(\mathcal{L}_h u_h))^{-1} - \lambda(u_h)^{-1})w_h\big)\big\|_{V_h} \\ &\le h^{-2}\big\|(I_h\lambda(\mathcal{L}_h u_h))^{-1} - \lambda(u_h)^{-1}\big\|_{L^2(\Omega_h)}\|w_h\|_{L^\infty} \\ &\le h^{-2}\big\|I_h\lambda(\mathcal{L}_h u_h) - \lambda(u_h)\big\|_{L^2(\Omega_h)}\|\Delta_h w_h\|_{L^2(\Omega_h)},\end{aligned}$$

where we used in the last step Lemma 3.3, the identity

$$(I_h\lambda(\mathcal{L}_h u_h))^{-1} - \lambda(u_h)^{-1} = (I_h\lambda(\mathcal{L}_h u_h))^{-1}\big(\lambda(u_h) - (I_h\lambda(\mathcal{L}_h u_h))\big)\lambda(u_h)^{-1},$$

and the maximum norm estimate on u_h. Finally, Lemma 5.2 leads to a uniform bound in h. Using the identity

$$Q_h\varphi - R_h\varphi = Q_h\big(\varphi - I_h\varphi\big) + R_h\big(I_h\varphi - \varphi\big),$$

the assertion follows, once we have established

$$h^{-2}\,\big\|\big(\mathrm{Id} - I_h\big)\big(\lambda(u_h)^{-1}w_h\big)\big\|_{L^2(\Omega_h)} + h^{-1}\,\big\|\big(\mathrm{Id} - I_h\big)\big(\lambda(u_h)^{-1}w_h\big)\big\|_{H_0^1(\Omega_h)} \le C.$$

However, applying Lemma 5.2 once more gives precisely this estimate. The bound in (5.2a) is derived in the very same way. □

Acknowledgments

The author thanks Selina Burkhard and Constantin Carle for their careful reading of this manuscript. Further, we like to thank Malik Scheifinger for providing us the codes for the numerical experiments.

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] D. Arndt, W. Bangerth, M. Feder, M. Fehling, R. Gassmöller, T. Heister, L. Heltai, M. Kronbichler, M. Maier, P. Munch, J.-P. Pelteret, S. Sticko, B. Turcksin, and D. Wells, *The deal.II library, version 9.4*, J. Numer. Math. **30** (2022), no. 3, 231–246.

[3] L. A. Bales, *Higher-order single-step fully discrete approximations for nonlinear second-order hyperbolic equations*, 1986, pp. 581–604. Hyperbolic partial differential equations, III. MR841989

[4] ———, *Error estimates for single step fully discrete approximations for nonlinear second order hyperbolic equations*, 1988, pp. 535–553. Hyperbolic partial differential equations. V. MR953564

[5] 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

[6] W. Bangerth, R. Hartmann, and G. Kanschat, *deal.II—a general-purpose object-oriented finite element library*, ACM Trans. Math. Software **33** (2007), no. 4, Art. 24, 27. MR2404402

[7] 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

[8] 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)

[9] P. G. Ciarlet, *The finite element method for elliptic problems*, Studies in Mathematics and its Applications, Vol. 4, North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978. MR0520174

[10] M. G. Crandall and P. E. Souganidis, *Convergence of difference approximations of quasilinear evolution equations*, Nonlinear Anal. **10** (1986), no. 5, 425–445. MR839356

[11] W. Dörfler, H. Gerner, and R. Schnaubelt, *Local well-posedness of a quasilinear wave equation*, Appl. Anal. **95** (2016), no. 9, 2110–2123. MR3515098

[12] B. Dörich and M. Hochbruck, *Exponential integrators for quasilinear wave-type equations*, SIAM J. Numer. Anal. **60** (2022), no. 3, 1472–1493. MR4442442

[13] B. Dörich, J. Leibold, and B. Maier, *Maximum norm error bounds for the full discretization of non-autonomous wave equations*, Technical Report 2021/47, Karlsruhe Institute of Technology, 2021.

[14] ———, *Optimal $W^{1,\infty}$-estimates for an isoparametric finite element discretization of elliptic boundary value problems*, Electron. Trans. Numer. Anal. **58** (2023), 1–21. MR4516523

[15] 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

[16] ———, *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

[17] R. E. Ewing, *On efficient time-stepping methods for nonlinear partial differential equations*, Comput. Math. Appl. **6** (1980), no. 1, Issu, 1–13. MR604081

[18] H. Fujita, N. Saito, and T. Suzuki, *Operator theory and numerical methods*, Studies in Mathematics and its Applications, vol. 30, North-Holland Publishing Co., Amsterdam, 2001. MR1854280

[19] L. Gauckler, J. Lu, J. L. Marzuola, F. Rousset, and K. Schratz, *Trigonometric integrators for quasilinear wave equations*, Math. Comp. **88** (2019), no. 316, 717–749. MR3882282

[20] 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

[21] M. Hochbruck and B. Maier, *Error analysis for space discretizations of quasilinear wave-type equations*, IMA J. Numer. Anal. **42** (2022), no. 3, 1963–1990. MR4454914

[22] 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

[23] M. Hochbruck, T. Pažur, and R. Schnaubelt, *Error analysis of implicit Runge-Kutta methods for quasilinear hyperbolic evolution equations*, Numer. Math. **138** (2018), no. 3, 557–579. MR3767694

[24] S. Kanda, *Convergence of difference approximations and nonlinear semigroups*, Proc. Amer. Math. Soc. **108** (1990), no. 3, 741–748. MR990425 (90h:47108)

[25] T. Kato, *Quasi-linear equations of evolution, with applications to partial differential equations*, Spectral theory and differential equations (Proc. Sympos., Dundee, 1974; dedicated to Konrad Jörgens), 1975, pp. 25–70. Lecture Notes in Math., Vol. 448. MR0407477

[26] ———, *Abstract differential equations and nonlinear mixed problems*, Lezioni Fermiane. [Fermi Lectures], Scuola Normale Superiore, Pisa; Accademia Nazionale dei Lincei, Rome, 1985. MR930267

[27] Y. Kobayashi, *Difference approximation of Cauchy problems for quasi-dissipative operators and generation of nonlinear semigroups*, J. Math. Soc. Japan **27** (1975), no. 4, 640–665. MR0399974 (53 #3812)

[28] B. Kovács and Ch. Lubich, *Stability and convergence of time discretizations of quasi-linear evolution equations of Kato type*, Numer. Math. **138** (2018), no. 2, 365–388. MR3748307

[29] J. Leibold, *A unified error analysis for the numerical solution of nonlinear wave-type equations with application to kinetic boundary conditions*, Ph.D. Thesis, 2021. https://doi.org/10.5445/IR/1000130222.

[30] B. Maier, *Error analysis for space and time discretizations of quasilinear wave-type equations*, Ph.D. Thesis, 2020.

[31] ———, *Error analysis for full discretizations of quasilinear wave-type equations with two variants of the implicit midpoint rule*, IMA J. Numer. Anal. **43** (2023), no. 2, 1149–1180. MR4568443

[32] C. G. Makridakis, *Finite element approximations of nonlinear elastic waves*, Math. Comp. **61** (1993), no. 204, 569–594. MR1195426

[33] M. Meliani and V. Nikolić, *Mixed approximation of nonlinear acoustic equations: Well-posedness and a priori error analysis*, arXiv, 2022. URL https://arxiv.org/abs/2209.02737.

[34] V. Nikolić and B. Wohlmuth, *A priori error estimates for the finite element approximation of Westervelt's quasi-linear acoustic wave equation*, SIAM J. Numer. Anal. **57** (2019), no. 4, 1897–1918. MR3987529

[35] J. A. Nitsche, *L_∞-convergence of finite element approximation*, Journées "Éléments Finis" (Rennes, 1975), 1975, pp. 18. MR568857

[36] T. Suzuki and H. Fujita, *A remark on the L^∞ bounds of the Ritz operator associated with a finite element approximation*, Numer. Math. **49** (1986), no. 5, 529–544. MR856383

[37] T. Takahashi, *Convergence of difference approximation of nonlinear evolution equations and generation of semigroups*, J. Math. Soc. Japan **28** (1976), no. 1, 96–113. MR399978

Appendix A. Modified Ritz map

In this section, we discuss the approximation property of $\widetilde{R}_h$ as well as its computation. Note that the same reasoning is valid in the conforming case.

Proof of Proposition 2.4. We use the definitions in (2.4b) and (2.11) to compute

$$\begin{aligned}\big|\big(\mathcal{L}_h^{V*}u - \widetilde{R}_h u \mid \varphi_h\big)_{V_h}\big| &= \big|\big(u \mid \mathcal{L}_h\varphi_h\big)_V - \big(\widetilde{I}_h u \mid \widetilde{\mathcal{L}_h}^{-1}\mathcal{L}_h\varphi_h\big)_{\widetilde{V}_h}\big| \\ &\le \big|\big(u - \widetilde{\mathcal{L}_h}\widetilde{I}_h u \mid \mathcal{L}_h\varphi_h\big)_V\big| \\ &\qquad + \big|\big(\widetilde{\mathcal{L}_h}\widetilde{I}_h u \mid \mathcal{L}_h\varphi_h\big)_V - \big(\widetilde{I}_h u \mid \widetilde{\mathcal{L}_h}^{-1}\mathcal{L}_h\varphi_h\big)_{\widetilde{V}_h}\big| \\ &= \Delta_1 + \Delta_2.\end{aligned}$$

We employ the stability of the lift in (2.2) and the interpolation property in (2.22) to obtain

$$\Delta_1 \lesssim h^{k'} \|u\|_{H^{k'+1}(\Omega)} \|\varphi_h\|_{V_h}.$$

The geometric estimate in [16, Lem. 8.24] together with (2.2) allows us to bound

$$\Delta_2 \lesssim h^{k'} \big\|\widetilde{I}_h u\big\|_{\widetilde{V}_h} \big\|\widetilde{\mathcal{L}_h}^{-1} \mathcal{L}_h \varphi_h\big\|_{V_h} \lesssim h^{k'} \|u\|_{H^{k'+1}(\Omega)} \|\varphi_h\|_{V_h},$$

and the claim follows setting $k' = k + 1$. Further, we use the definition of Δ_h in (2.9) and the inverse estimate (2.25)

$$\|\Delta_h u_h\|_{L^2}^2 = -\left(u_h \mid \Delta_h u_h\right)_{V_h} \lesssim h^{-1} \|u_h\|_{V_h} \|\Delta_h u_h\|_{L^2}, \tag{A.1}$$

and obtain the second bound with one power less in h. □

In order to compute (2.11), we have to solve a linear system with the stiffness matrix corresponding to the bilinear form $(\cdot \mid \cdot)_{V_h}$ and right-hand side $\widetilde{\ell}_u$. For a basis φ_i, $i = 1, \dots, L$, of V_h, the entries are given by

$$(\widetilde{\ell}_u)_i = \big(\widetilde{I}_h u \mid \widetilde{\mathcal{L}_h}^{-1} \mathcal{L}_h \varphi_i\big)_{\widetilde{V}_h}.$$

Going through the construction explained in [15, Sec. 4.1.2–4.2], one observes that $\widetilde{\varphi}_i \coloneqq \widetilde{\mathcal{L}_h}^{-1} \mathcal{L}_h \varphi_i \in \widetilde{V}_h$, however $\widetilde{\varphi}_i$ is not a nodal basis function. Since also $\widetilde{I}_h u \in \widetilde{V}_h$, one only needs to modify the routines which are used to assemble the stiffness matrix corresponding to $(\cdot \mid \cdot)_{\widetilde{V}_h}$. In particular, denoting for the reference element $\widehat{K}$ by $\widehat{\varphi}_i$ and $\widehat{\psi}_j$ the nodal basis polynomials of $\mathcal{P}^k(\widehat{K})$ and $\mathcal{P}^{k'}(\widehat{K})$, respectively, one is left to compute the inner products

$$\big(\nabla \widehat{\psi}_j \mid \nabla \widehat{\varphi}_i\big)_{L^2(\widehat{K})}.$$

Then, the transformation maps to the elements in $\widetilde{\Omega}_h$ are used to assemble the right-hand side $\widetilde{\ell}_u$.

Appendix B. Interpolation and adjoint lift

In this appendix, we provide the proof of Lemma 2.10, (2.23), and Lemma 5.2. The following estimate appears to be standard, but since we could not find a reference in the literature, we provide its proof here.

Lemma B.1. *For $m = 0, 1$, there is a constant $C_m > 0$ independent of h such that*

$$\|I_h(\mathcal{L}_h \varphi_h \cdot \mathcal{L}_h \psi_h)\|_{W^{m,2}(\Omega_h)} \le C_m \|\mathcal{L}_h \varphi_h \cdot \mathcal{L}_h \psi_h\|_{W^{m,2}(\Omega_h)},$$

for all $\varphi_h, \psi_h \in V_h$.

Proof. Writing $\mathcal{L}_h I_h = \mathrm{Id} + (\mathcal{L}_h I_h - \mathrm{Id})$, by (2.2) it is sufficient to show the assertion for $\mathcal{L}_h I_h - \mathrm{Id}$ instead of I_h. Passing to the reference cell $\widehat{K}$, we only consider the case $m = 1$. We define the map

$$\mathrm{Id} - \widehat{I}_h \colon \big(\mathcal{P}_{2k}(\widehat{K}), |\cdot|_{H^1}\big) \to \big(\mathcal{P}_k(\widehat{K}), \|\cdot\|_{H^1}\big),$$

which is a bounded linear operator with a constant C, such that for any $\varphi \in \mathcal{P}_{2k}(\widehat{K})$

$$\big\|(\mathrm{Id} - \widehat{I}_h)\varphi\big\|_{H^1(\widehat{K})} \le C\, |\varphi|_{H^1(\widehat{K})}.$$

Then, employing [16, Lem. 4.12] yields the result on an arbitrary cell K. □

With this, we directly conclude the desired stability estimate.

Proof of Lemma 2.10. By the nodal interpolation property (2.3) and Lemma B.1, we obtain

$$\|I_h(\varphi \cdot \mathcal{L}_h\psi_h)\|_{V_h} = \|I_h(\mathcal{L}_h I_h^e\varphi \cdot \mathcal{L}_h\psi_h)\|_{V_h} \leq C\, \|\mathcal{L}_h I_h^e \varphi \cdot \mathcal{L}_h\varphi_h\|_{V_h}.$$

Using Sobolev's embedding and the stability of the interpolation from (2.22), we further estimate

$$\|\mathcal{L}_h I_h^e\varphi \cdot \mathcal{L}_h\varphi_h\|_{V_h} \leq C\, \|\mathcal{L}_h I_h^e\varphi\|_{W^{1,N+\delta}} \|\varphi_h\|_{V_h} \leq C\, \|\varphi\|_{W^{1,N+\delta}} \|\varphi_h\|_{V_h}.$$

By the same reasoning, we obtain the bound in the L^2-norm. □

Proof of (2.23). The stability of $\mathcal{L}_h^{H*}$ directly follows from the definition (2.4) and the stability of the lift $\mathcal{L}_h$ in (2.2). For the error bound, we first observe

$$\begin{aligned}\left\|(I_h - \mathcal{L}_h^{H*})\varphi\right\|_{H_h} &= \sup_{\|\psi_h\|_{H_h}=1} \big((I_h - \mathcal{L}_h^{H*})\varphi \mid \psi_h\big)_{L^2(\Omega_h)} \\ &= \sup_{\|\psi_h\|_{H_h}=1} \Big(\big((\mathcal{L}_h I_h - \mathrm{Id})\varphi \mid \mathcal{L}_h\psi_h\big)_{L^2(\Omega)} \\ &\qquad + \big(I_h\varphi \mid \psi_h\big)_{L^2(\Omega_h)} - \big(\mathcal{L}_h I_h\varphi \mid \mathcal{L}_h\psi_h\big)_{L^2(\Omega)}\Big).\end{aligned}$$

For the first term we apply (2.22), and for the difference we use [16, Lem. 8.24] to obtain

$$\left\|(I_h - \mathcal{L}_h^{H*})\varphi\right\|_{H_h} \lesssim h^{\ell+1}\, \|\varphi\|_{H^{\ell+1}(\Omega)} + h^{\ell}\, \|\mathcal{L}_h I_h\varphi\|_{L^2(U_h)}$$

with the boundary layer $U_h \coloneqq \{x \in \Omega \mid \mathrm{dist}(x, \partial\Omega) \leq h\}$. Below, we show

$$\|\mathcal{L}_h I_h\varphi\|_{L^\infty(U_h)} \lesssim h\, \|I_h\varphi\|_{W^{1,\infty}(\Omega_h)} \tag{B.1}$$

and use this together with $\mathrm{vol}(U_h)^{1/2} \lesssim h^{1/2}$ to estimate

$$\|\mathcal{L}_h I_h\varphi\|_{L^2(U_h)} \lesssim h^{1/2}\, \|\mathcal{L}_h I_h\varphi\|_{L^\infty(U_h)} \lesssim h^{3/2}\, \|I_h\varphi\|_{W^{1,\infty}(\Omega_h)} \lesssim h\, \|I_h\varphi\|_{W^{1,6}(\Omega_h)},$$

where we used the inverse inequality (2.25) in the last step. The stability of the interpolation (2.22) and the Sobolev embedding yield

$$\|\mathcal{L}_h I_h\varphi\|_{L^2(U_h)} \lesssim h\, \|\varphi\|_{H^2(\Omega)}, \tag{B.2}$$

and thus the assertion. To show (B.1), we pick some $x_0 \in U_h$ and $y_0 \in \partial\Omega$ with $|x_0 - y_0| \leq h$ such that

$$(\mathcal{L}_h I_h\varphi)(y_0) = 0, \quad |(\mathcal{L}_h I_h\varphi)(x_0)| = \|\mathcal{L}_h I_h\varphi\|_{L^\infty(U_h)}.$$

Then we use the fundamental theorem of calculus to see

$$\begin{aligned}|(\mathcal{L}_h I_h\varphi)(x_0)| &= |(\mathcal{L}_h I_h\varphi)(x_0) - (\mathcal{L}_h I_h\varphi)(y_0)| \\ &= \Big|\int_0^1 (\nabla\mathcal{L}_h I_h\varphi)(sx_0 + (1-s)y_0)(x_0 - y_0)\, ds\Big| \\ &\leq \|\mathcal{L}_h I_h\varphi\|_{W^{1,\infty}(\Omega)}\, |x_0 - y_0|,\end{aligned}$$

which gives the assertion. □

For the interpolation estimate in Lemma 5.2, we exploit that the $(k+1)$-st derivative vanishes for polynomials of degree k. With this one can gain an additional power of h, but does not have to apply an inverse estimate to the highest derivative. In a different context, this was also used by Nitsche in [35, p. 7].

Proof of Lemma 5.2. We perform the proof in the case $L = 2$, and explain the generalization in the end. Thus, consider $\varphi_h, \psi_h \in V_h$ satisfying the assumptions of the lemma. We expand the expression over all elements, which gives

$$\begin{aligned}\|f(\varphi_h,\psi_h) - I_h f(\mathcal{L}_h\varphi_h, \mathcal{L}_h\psi_h)\|^2_{L^2(\Omega_h)} &= \sum_K \|f(\varphi_h,\psi_h) - I_h f(\mathcal{L}_h\varphi_h, \mathcal{L}_h\psi_h)\|^2_{L^2(K)} \\ &\lesssim h^{2(k+1)} \sum_K |f(\varphi_h,\psi_h)|^2_{H^{k+1}(K)} \\ &\lesssim h^4 \sum_K h^{2(k-1)} |f(\varphi_h,\psi_h)|^2_{H^{k+1}(K)},\end{aligned}$$

as well as

$$\|f(\varphi_h,\psi_h) - I_h f(\mathcal{L}_h\varphi_h, \mathcal{L}_h\psi_h)\|^2_{H^1(\Omega_h)} \lesssim h^2 \sum_K h^{2(k-1)} |f(\varphi_h,\psi_h)|^2_{H^{k+1}(K)}.$$

In the following, we show on each element K

$$\begin{aligned}h^{2(k-1)} |f(\varphi_h,\psi_h)|^2_{H^{k+1}(K)} \le C\big(&\|f(\varphi_h,\psi_h)\|^2_{L^2(K)} + \|\varphi_h\|^2_{H^1(K)} + \|\psi_h\|^2_{H^1(K)} \\ &+ \|\varphi_h\|^4_{W^{1,4}(K)} + \|\psi_h\|^4_{W^{1,4}(K)}\big),\end{aligned}$$

which, by summing over all K, gives the assertion by the Cauchy–Schwarz inequality

$$\begin{aligned}\sum_K \|\varphi_h\|_{H^1(K)} \|\psi_h\|_{H^1(K)} &\le \|\varphi_h\|_{H^1(\Omega_h)} \|\psi_h\|_{H^1(\Omega_h)}, \\ \sum_K \|\varphi_h\|^2_{W^{1,4}(K)} \|\psi_h\|^2_{W^{1,4}(K)} &\le \|\varphi_h\|^2_{W^{1,4}(\Omega_h)} \|\psi_h\|^2_{W^{1,4}(\Omega_h)}.\end{aligned}$$

The constant C depends on the maximum norm of φ_h and ψ_h and is thus uniformly bounded by assumption.

If we denote by $\widehat{K}$ the reference element, and denote by A_K the affine part of the element maps, we know from [16, Lem. 4.12] that

$$\begin{aligned}\|\widehat{\varphi}_h\|_{L^2(\widehat{K})} &\lesssim |\det A_K|^{-1/2} \|\varphi_h\|_{L^2(K)}, \\ |\widehat{\varphi}_h|_{H^1(\widehat{K})} &\lesssim h|\det A_K|^{-1/2} \|\varphi_h\|_{H^1(K)}, \\ |\widehat{\varphi}_h|^2_{W^{1,4}(\widehat{K})} &\lesssim h^2|\det A_K|^{-1/2} \|\varphi_h\|^2_{W^{1,4}(K)}.\end{aligned}$$

On the other hand, by the same lemma [16, Lem. 4.12] it holds

$$h^{k-1} |f(\varphi_h,\psi_h)|_{H^{k+1}(K)} \lesssim |\det A_K|^{1/2} \sum_{r=0}^{k+1} h^{k-(r+1)} \big|f(\widehat{\varphi}_h, \widehat{\psi}_h)\big|_{H^r(\widehat{K})}.$$

We now treat the summands separately, and show

(B.3)
$$\begin{aligned}&|\det A_K|^{1/2} h^{k-(r+1)} \big|f(\widehat{\varphi}_h, \widehat{\psi}_h)\big|_{H^r(\widehat{K})} \\ &\lesssim \big(\|f(\varphi_h,\psi_h)\|^2_{L^2(K)} + \|\varphi_h\|^2_{H^1(K)} + \|\psi_h\|^2_{H^1(K)} + \|\varphi_h\|^4_{W^{1,4}(K)} + \|\psi_h\|^4_{W^{1,4}(K)}\big)\end{aligned}$$

for $r = 0, \dots, k+1$, which then implies our claim. In the case $r = 0$, we directly obtain

$$|\det A_K|^{1/2} h^{k-1} \big\|f(\widehat{\varphi}_h, \widehat{\psi}_h)\big\|_{L^2(\widehat{K})} \lesssim \|f(\varphi_h,\psi_h)\|_{L^2(K)},$$

and (B.3) follows.

For $r \geq 1$, we use the inverse estimate on the reference element in the forms

$$\|\nabla^m \widehat{\varphi}_h\|_{L^q(\widehat{K})} \lesssim \|\nabla \widehat{\varphi}_h\|_{L^q(\widehat{K})}, \quad m \geq 1, q \in [1,\infty],$$
$$\|\nabla \widehat{\varphi}_h\|_{L^q(\widehat{K})} \lesssim \|\widehat{\varphi}_h\|_{L^\infty(\widehat{K})}, \quad q \in [1,\infty],$$

differentiate the term $f(\widehat{\varphi}_h, \widehat{\psi}_h)$, and reduce all derivatives by the inverse estimate to $\nabla \widehat{\varphi}_h$. Then using the maximum norm bound on $\widehat{\varphi}_h$ and $\widehat{\psi}_h$, we obtain

$$\big|f(\widehat{\varphi}_h, \widehat{\psi}_h)\big|_{H^r(\widehat{K})} \lesssim \Big(\big\|\nabla^r \widehat{\varphi}_h\big\|_{L^2} + \big(\big\|\nabla \widehat{\varphi}_h\big\|_{L^4} + \big\|\nabla \widehat{\psi}_h\big\|_{L^4}\big)^2 + \big\|\nabla^r \widehat{\psi}_h\big\|_{L^2}\Big). \tag{B.4}$$

Note that for $r = 1$, the quadratic term can be dropped. For $r \leq k$, we employ the inverse estimate once more, to obtain with $k-(r+1) \geq -1$

$$\begin{aligned} h^{k-(r+1)}\big|f(\widehat{\varphi}_h, \widehat{\psi}_h)\big|_{H^r(\widehat{K})} &\lesssim h^{-1}\big(\big\|\nabla \widehat{\varphi}_h\big\|_{L^2} + \big\|\nabla \widehat{\psi}_h\big\|_{L^2}\big) \\ &\lesssim |\det A_K|^{-1/2}\big(\|\varphi_h\|_{H^1(K)} + \|\psi_h\|_{H^1(K)}\big), \end{aligned}$$

and (B.3) also follows for $1 \leq r \leq k$.

For $r = k+1$ we exploit that $\partial^{k+1}\widehat{\varphi}_h = 0$, and thus (B.4) yields

$$\begin{aligned} h^{-2}\,|f(\widehat{\varphi}_h)|_{H^{k+1}(\widehat{K})} &\lesssim h^{-2}\big(\big\|\nabla \widehat{\varphi}_h\big\|_{L^4} + \big\|\nabla \widehat{\psi}_h\big\|_{L^4}\big)^2 \\ &\lesssim |\det A_K|^{-1/2}\big(\|\varphi_h\|^2_{W^{1,4}(K)} + \|\psi_h\|^2_{W^{1,4}(K)}\big). \end{aligned}$$

This gives the claim of the lemma in the case $L = 2$.

In order to treat $L > 2$, we only need a modification of (B.4) since nothing changes for $r = 0$. A straightforward computation gives

$$\big|f(\widehat{\varphi}_{1,h}, \ldots \widehat{\varphi}_{L,h})\big|_{H^r(\widehat{K})} \lesssim \sum_{i=1}^{L} \big\|\nabla^r \widehat{\varphi}_{i,h}\big\|_{L^2} + \sum_{i,j=1}^{L} \big\|\nabla \widehat{\varphi}_{i,h}\big\|_{L^4}\big\|\nabla \widehat{\varphi}_{j,h}\big\|_{L^4},$$

and the same ideas apply. □

Appendix C. Discrete Sobolev embedding

The proof is adapted from the conforming case presented in [7, Lem. 4.1], but is able to cover a larger range of exponents. Similar results including the discrete differential operator Δ_h are shown in [18, Thm. 1.12] and [36, Thm. 3].

Proof of Lemma 3.3. First, we define the inverse S_h of Δ_h form (2.9) by

$$(S_h \varphi_h \mid \psi_h)_{V_h} = -(\varphi_h \mid \psi_h)_{H_h}, \quad \varphi_h, \psi_h \in V_h,$$

and its continuous counterpart $S = \Delta^{-1}$ satisfying

$$(S\varphi \mid \psi)_V = -(\varphi \mid \psi)_H, \quad \varphi, \psi \in V.$$

We further define the modified solution operator $\widetilde{S}_h = \mathcal{L}_h^{V*} S \mathcal{L}_h$, and write $S_h = \widetilde{S}_h + (S_h - \widetilde{S}_h)$. For the first term, we use the stability of the Ritz map in W^{1,p^*} from (2.24) with $\ell = 0$ and (2.14) to obtain

$$\big\|\widetilde{S}_h \varphi_h\big\|_{W^{1,p^*}(\Omega_h)} \lesssim \big\|S\mathcal{L}_h \varphi_h\big\|_{W^{1,p^*}(\Omega)} \lesssim \big\|S\mathcal{L}_h \varphi_h\big\|_{H^2(\Omega)} \lesssim \big\|\varphi_h\big\|_{L^2(\Omega_h)}.$$

It remains to bound the difference, stemming from the nonconformity, by the inverse estimate (2.25)

$$\begin{aligned}\big\|\widetilde{S}_h\varphi_h - S_h\varphi_h\big\|_{W^{1,p^*}(\Omega_h)} &\le Ch^{N/p^*-N/2}\big\|\widetilde{S}_h\varphi_h - S_h\varphi_h\big\|_{V_h}\\ &\le Ch^{-1}\sup_{\|\psi_h\|_{V_h}=1}\Big(\widetilde{S}_h\varphi_h - S_h\varphi_h \mid \psi_h\Big)_{V_h}\\ &= Ch^{-1}\sup_{\|\psi_h\|_{V_h}=1}\big((\varphi_h \mid \psi_h)_{H_h} - (\mathcal{L}_h\varphi_h \mid \mathcal{L}_h\psi_h)_H\big).\end{aligned}$$

We use [16, Lem. 8.24] to obtain

$$\big|(\varphi_h \mid \psi_h)_{H_h} - (\mathcal{L}_h\varphi_h \mid \mathcal{L}_h\psi_h)_H\big| \lesssim h\,\|\varphi_h\|_{L^2}\,\|\psi_h\|_{V_h},$$

which yields

$$\|S_h\varphi_h\|_{L^\infty(\Omega_h)} + \|S_h\varphi_h\|_{W^{1,p^*}(\Omega_h)} \le C\,\|\varphi_h\|_{L^2(\Omega_h)},$$

and hence the assertion. □

Appendix D. Chain rule for the discrete differential operator

In this section, we provide the proof of Lemma 5.3. We recall the Ritz projection defined in (2.5) and first show a crucial bound in L^2.

Lemma D.1. *Let $u_h, \varphi_h \in V_h$ and assume that $\mu\colon \mathbb{R} \to \mathbb{R}$ is continuously differential. Then it holds*

$$\|R_h(\mu(u_h)\varphi_h)\|_{L^2(\Omega_h)} \le C\,\|\mu(u_h)\|_{W^{1,N+\delta}(\Omega_h)}\,\|\varphi_h\|_{L^2(\Omega_h)}$$

with a constant C independent of h.

Proof. We first estimate

$$\|R_h(\mu(u_h)\varphi_h)\|_{L^2(\Omega_h)} \le \|\mu(u_h)\varphi_h\|_{L^2(\Omega_h)} + \|(\mathrm{Id} - R_h)(\mu(u_h)\varphi_h)\|_{L^2(\Omega_h)},$$

and show in the following

$$\begin{aligned}\|(\mathrm{Id} - R_h)(\mu(u_h)\varphi_h)\|_{L^2(\Omega_h)} &\lesssim h\,\|(\mu(u_h)\varphi_h)\|_{H^1(\Omega_h)}\\ &\lesssim h\,\|\mu(u_h)\|_{W^{1,N+\delta}(\Omega_h)}\,\|\varphi_h\|_{H^1(\Omega_h)}.\end{aligned} \tag{D.1}$$

Using the inverse estimate (2.25), we conclude the assertion.

We now show (D.1) by an Aubin–Nitsche trick. We define $e = (\mathrm{Id} - R_h)w$, for some $w \in H_0^1(\Omega_h)$, and consider the solution $z \in H^2(\Omega) \cap H_0^1(\Omega)$ of

$$(z \mid \varphi)_{H_0^1(\Omega)} = (\mathcal{L}_h e \mid \varphi)_{L^2(\Omega)}, \quad \varphi \in H_0^1(\Omega).$$

This gives

$$\begin{aligned}\|\mathcal{L}_h e\|_{L^2(\Omega)}^2 &= (z \mid \mathcal{L}_h e)_{H_0^1(\Omega)}\\ &= (z - \mathcal{L}_h I_h z \mid \mathcal{L}_h e)_{H_0^1(\Omega)} + (\mathcal{L}_h I_h z \mid \mathcal{L}_h e)_{H_0^1(\Omega)} - (I_h z \mid e)_{H_0^1(\Omega_h)}\\ &\quad + (I_h z \mid e)_{H_0^1(\Omega_h)}.\end{aligned}$$

The first term is bounded using (2.22) and elliptic regularity by

$$(z - \mathcal{L}_h I_h z \mid \mathcal{L}_h e)_{H_0^1(\Omega)} \lesssim h\,\|z\|_{H^2}\,\|e\|_{H^1} \lesssim h\,\|e\|_{L^2}\,\|e\|_{H^1},$$

and, using the geometric estimates in [16, Lem. 8.24], for the second term it holds

$$(\mathcal{L}_h I_h z \mid \mathcal{L}_h e)_{H_0^1(\Omega)} - (I_h z \mid e)_{H_0^1(\Omega_h)} \lesssim h\,\|I_h z\|_{H^1}\,\|e\|_{H^1} \lesssim h\,\|e\|_{L^2}\,\|e\|_{H^1}.$$

By the definition of e and R_h, we have due to $I_h z \in V_h$

$$(I_h z \mid e)_{H_0^1(\Omega_h)} = (I_h z \mid w - R_h w)_{H_0^1(\Omega_h)} = 0,$$

and the claim follows. □

With this, we obtain the boundedness in the discrete chain rule.

Proof of Lemma 5.3. We use the definition of Δ_h in (2.9) to obtain

$$\begin{aligned}
\|\Delta_h R_h(\mu(u_h)w_h)\|_{L^2(\Omega_h)} &= \sup_{\|\varphi_h\|_{H_h}\le 1} (\mu(u_h)w_h \mid \varphi_h)_{H_0^1} \\
&= \sup_{\|\varphi_h\|_{H_h}\le 1} \int_{\Omega_h} \mu'(u_h)\nabla u_h w_h \nabla\varphi_h + \mu(u_h)\nabla w_h \nabla \varphi_h \, dx.
\end{aligned}$$

The idea now is to express the integral in terms of $\Delta_h u_h$ and $\Delta_h w_h$ and lower order terms, where the is no gradient on φ_h. Employing the identities

$$\begin{aligned}
\mu'(u_h)w_h\nabla\varphi_h &= \nabla\big(\mu'(u_h)w_h\varphi_h\big) - \nabla(\mu'(u_h)w_h)\varphi_h, \\
\mu(u_h)\nabla\varphi_h &= \nabla\big(\mu(u_h)\varphi_h\big) - \mu'(u_h)\nabla u_h\varphi_h,
\end{aligned}$$

we derive

$$\begin{aligned}
&\int_{\Omega_h} \mu'(u_h)\nabla u_h w_h \nabla\varphi_h + \mu(u_h)\nabla w_h\nabla\varphi_h \, dx \\
&= \int_{\Omega_h} \nabla u_h \nabla\big(\mu'(u_h)w_h\varphi_h\big)\, dx - \int_{\Omega_h} \nabla u_h\nabla(\mu'(u_h)w_h)\varphi_h \, dx \\
&\quad + \int_{\Omega_h} \nabla w_h\nabla\big(\mu(u_h)\varphi_h\big)\, dx - \int_{\Omega_h} \nabla w_h(\mu'(u_h)\nabla u_h)\varphi_h\, dx \\
&= (u_h \mid \mu'(u_h)w_h\varphi_h)_{H^1} + (w_h \mid \mu(u_h)\varphi_h)_{H^1} \\
&\quad - \int_{\Omega_h} \nabla u_h\nabla(\mu'(u_h)w_h)\varphi_h\, dx - \int_{\Omega_h} \nabla w_h(\mu'(u_h)\nabla u_h)\varphi_h\, dx \\
&= -(\Delta_h u_h \mid R_h\mu'(u_h)w_h\varphi_h)_{L^2} - (\Delta_h w_h \mid R_h\mu(u_h)\varphi_h)_{L^2} \\
&\quad - \int_{\Omega_h} \nabla u_h\nabla(\mu'(u_h)w_h)\varphi_h\, dx - \int_{\Omega_h} \nabla w_h(\mu'(u_h)\nabla u_h)\varphi_h\, dx.
\end{aligned}$$

This yields with the stability of R_h shown in Lemma D.1 the estimate

$$\begin{aligned}
\|\Delta_h R_h(\mu(u_h)w_h)\|_{L^2(\Omega_h)} &\lesssim \|\Delta_h u_h\|_{L^2} \|\mu'(u_h)w_h\|_{W^{1,N+\delta}} + \|\Delta_h w_h\|_{L^2} \|\mu(u_h)\|_{W^{1,N+\delta}} \\
&\quad + \|\nabla u_h\nabla(\mu'(u_h)w_h)\|_{L^2} + \|\nabla w_h\nabla(\mu'(u_h)u_h)\|_{L^2}.
\end{aligned}$$

We then use Lemma 3.3 in the form $\|\psi_h\|_{W^{1,N+\delta}} \lesssim \|\Delta_h\psi_h\|_{L^2}$, and thus the first two terms are bounded. The other two are similar to each other, and it thus suffices to bound

$$\begin{aligned}
\|\nabla u_h\nabla(\mu'(u_h)w_h)\|_{L^2} &\le \|\nabla u_h\mu''(u_h)\nabla u_h w_h\|_{L^2} + \|\nabla u_h\mu'(u_h)\nabla w_h\|_{L^2} \\
&\lesssim \|\nabla u_h\|_{L^4}^2 + \|\nabla u_h\|_{L^4}\|\nabla w_h\|_{L^4},
\end{aligned}$$

where we again used the maximum norm bounds u_h and w_h and $\mu \in C^2(\mathbb{R})$, and the claim is shown. □

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

Email address: `benjamin.doerich@kit.edu`