Karlsruhe Institute of Technology

Wave phenomena
analysis and numerics

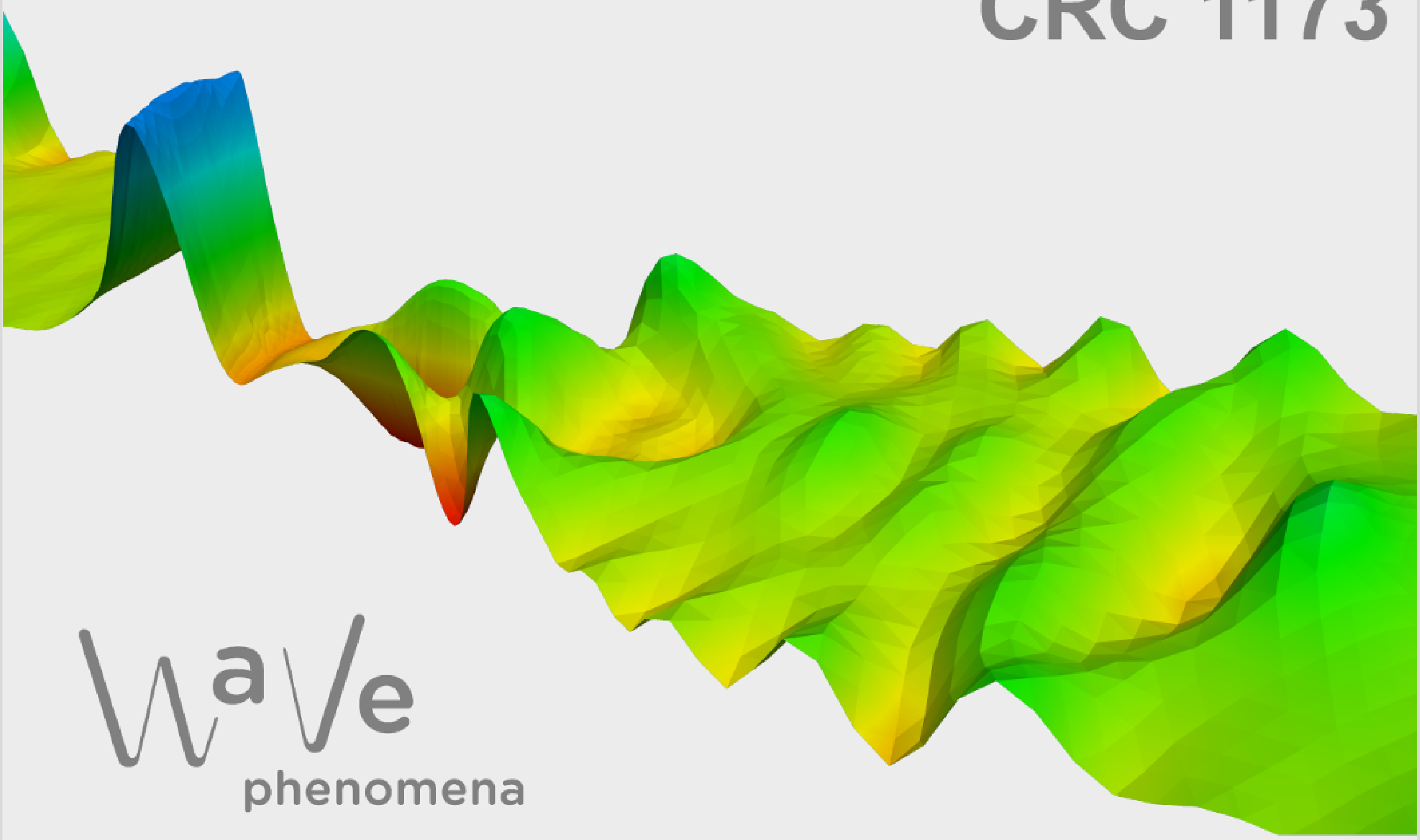
Space-time discontinuous Galerkin methods for weak solutions of hyperbolic linear symmetric Friedrichs systems

Daniele Corallo, Willy Dörfler, Christian Wieners

CRC Preprint 2022/36, August 2022

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173



Wave phenomena

Participating universities

Universität Stuttgart

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

TU WIEN

Funded by

DFG

ISSN 2365-662X

SPACE-TIME DISCONTINUOUS GALERKIN METHODS FOR WEAK SOLUTIONS OF HYPERBOLIC LINEAR SYMMETRIC FRIEDRICHS SYSTEMS

Daniele Corallo[1], Willy Dörfler[2] and Christian Wieners[*,3]

Abstract. We study weak solutions and its approximation of hyperbolic linear symmetric Friedrichs systems describing acoustic, elastic, or electro-magnetic waves. For the corresponding first-order systems we construct discontinuous Galerkin discretizations in space and time with full upwind, and we show primal and dual consistency. Stability and convergence estimates are provided with respect to a mesh-dependent DG norm which includes the L_2 norm at final time. Numerical experiments confirm that the a piori results are of optimal order also for solutions with low regularity, and we show that the error in the DG norm can be closely approximated with a residual-type error indicator.

1991 Mathematics Subject Classification 65N30.

August 23, 2022.

1. Introduction

Linear wave equations are hyperbolic, and the formulation as first-order symmetric Friedrichs system provides a well established setting for analyzing and approximating solutions. A specific feature of hyperbolic systems is the transport of discontinuities along characteristics. Our goal is to provide a numerical scheme which is efficient for smooth solutions as well as for weak solutions with discontinuities.

For smooth solutions of linear symmetric Friedrichs systems $\mathcal{O}(h^{s-1/2})$ convergence can be established for discontinuous Galerkin approximations in space with respect to suitable mesh-dependent DG norm [Ern and Guermond, 2021, Chap. 57], [Di Pietro and Ern, 2011, Chap. 7]. For acoustics, the convergence analysis of a space-time approximation in a DG semi-norm provides estimates for all discrete time steps [Bansal et al., 2021, Prop. 6.5].

Finite volume convergence $\mathcal{O}(h^{1/2})$ for hyperbolic linear symmetric Friedrichs systems is established in [Jovanović and Rohde, 2005] combined with first-order time-stepping. Discontinuous Galerkin methods in time are analyzed in [Falk and Richter, 1999] for tent-type space-time meshes. This is adapted to space-time discontinuous Galerkin methods on general space-time meshes with upwind flux for acoustics in [Bansal et al., 2021], where the convergence is established for sufficiently smooth solutions based on estimates in a suitable DG semi-norm. In particular, the analysis includes the adaptive approximation of corner singularities.

Here, we consider a DG method in space and time for linear symmetric Friedrichs systems, and we show inf-sup stability and convergence in the DG norm. Therefore we transfer our results for space-time Petrov–Galerkin methods in [Dörfler et al., 2016, Dörfler et al., 2019] with continuous approximations in time and for the DPG method in [Ernesti and Wieners, 2019a, Ernesti and Wieners, 2019b], where convergence in a stronger graph norm is considered. Our analysis includes bounds for the consistency error in the case that piecewise discontinuous material parameters are not aligned with the mesh. Convergence in the limit for piecewise discontinuous solutions of Riemann problems is established only in L_2.

The space-time method is realized in the parallel finite element system M++ [Baumgarten and Wieners, 2021]. In our numerical examples we confirm the a priori estimates for weak as well as for smooth solutions, and we demonstrate the efficiency of the p-adaptive scheme.

The paper is organized as follows. In Sect. 2 we introduce the notation and the formulation of wave equations as first-order systems, in Sect. 3 we introduce the DG discretization in time and in space. In Sect. 4 we consider well-posedness and stability, in Sect. 5 we prove existence of weak solutions and convergence estimates, in Sect. 5.3 we introduce an a

Keywords and phrases: weak solution of linear symmetric Friedrichs systems, discontinuous Galerkin methods in space and time, error estimators for first-order systems

[1] Institute of Applied and Numerical Mathematics, KIT, Karlsruhe, Germany. Email: `daniele.corallo@kit.edu`

[2] Institute of Applied and Numerical Mathematics, KIT, Karlsruhe, Germany. Email: `willy.doerfler@kit.edu`

[3] Institute of Applied and Numerical Mathematics, KIT, Karlsruhe, Germany. Email: `christian.wieners@kit.edu` (corresponding author)

posteriori error indicator, and in Sect. 6 we present numerical results. In Sect. 7 we conclude with a discussion of possible extensions and open problems.

2. Symmetric Friedrichs systems

We consider weak solutions of linear hyperbolic first-order systems in the form of symmetric Friedrichs systems. Let $\Omega \subset \mathbb{R}^d$ be a bounded domain in space with Lipschitz boundary $\partial\Omega$, $I = (0,T)$ a time interval, and we denote the space-time cylinder by $Q = (0,T)\times\Omega$. Boundary conditions will be imposed on $\Gamma_k \subset \partial\Omega$ for $k = 1,\dots,m$ depending on the model, where m is the dimension of the first-order system.

For $S \subset Q$ the L_2 norm and inner product are denoted by $\|\cdot\|_S$ and $(\cdot,\cdot)_S$.

Let $L = M\partial_t + A$ be a linear differential operator in space and time, where $(M\mathbf{v})(t,\mathbf{x}) = \underline{M}(\mathbf{x})\mathbf{v}(t,\mathbf{x})$ defines the operator M with a uniformly positive definite matrix-valued function $\underline{M} \in \mathrm{L}_\infty(\Omega;\mathbb{R}^{m\times m}_{\mathrm{sym}})$, and where $A\mathbf{v} = \sum_{j=1}^d \underline{A}_j\partial_j\mathbf{v}$ is a differential operator in space with matrices $\underline{A}_j \in \mathbb{R}^{m\times m}_{\mathrm{sym}}$. Since $\underline{M}$ is uniformly positive definite, constants $C_M \ge c_M > 0$ exists such that

$$c_M\,\mathbf{y}^\top\mathbf{y} \le \mathbf{y}^\top\underline{M}(\mathbf{x})\mathbf{y} \le C_M\,\mathbf{y}^\top\mathbf{y}\,, \qquad \mathbf{y}\in\mathbb{R}^m \text{ and a.a. } \mathbf{x}\in\Omega\,.$$

We observe

$$\big(L\mathbf{v},\mathbf{w}\big)_Q = \big(M\partial_t\mathbf{v},\mathbf{w}\big)_Q + \big(A\mathbf{v},\mathbf{w}\big)_Q = -\big(\mathbf{v},M\partial_t\mathbf{w}\big)_Q - \big(\mathbf{v},A\mathbf{w}\big)_Q = -\big(\mathbf{v},L\mathbf{w}\big)_Q\,, \qquad \mathbf{v},\mathbf{w}\in\mathrm{C}^1_{\mathrm{c}}(Q;\mathbb{R}^m)\,,$$

so that $L^* = -L$ is the adjoint differential operator. This is now complemented by initial and boundary conditions.

For the unit normal vector $\mathbf{n}\in\mathrm{L}_\infty(\partial\Omega;\mathbb{R}^d)$ we define the matrix $\underline{A}_{\mathbf{n}} = \sum_{j=1}^d n_j\underline{A}_j \in \mathbb{R}^{m\times m}_{\mathrm{sym}}$, so that

$$\big(A\mathbf{v},\mathbf{w}\big)_\Omega + \big(\mathbf{v},A\mathbf{w}\big)_\Omega = \big(\underline{A}_{\mathbf{n}}\mathbf{v},\mathbf{w}\big)_{\partial\Omega} = \big(\mathbf{v},\underline{A}_{\mathbf{n}}\mathbf{w}\big)_{\partial\Omega}\,, \qquad \mathbf{v},\mathbf{w}\in\mathrm{C}^1(\overline{\Omega};\mathbb{R}^m)\,.$$

Correspondingly, we get for the operator L in space and time

$$\big(L\mathbf{v},\mathbf{w}\big)_Q + \big(\mathbf{v},L\mathbf{w}\big)_Q = \big(M\mathbf{v}(T),\mathbf{w}(T)\big)_\Omega - \big(M\mathbf{v}(0),\mathbf{w}(0)\big)_\Omega + \big(\underline{A}_{\mathbf{n}}\mathbf{v},\mathbf{w}\big)_{(0,T)\times\partial\Omega}\,, \qquad \mathbf{v},\mathbf{w}\in\mathrm{C}^1(\overline{Q};\mathbb{R}^m)\,,$$

i.e., inserting $L^* = -L$,

$$\big(\mathbf{v},L^*\mathbf{w}\big)_Q = \big(L\mathbf{v},\mathbf{w}\big)_Q - \big(M\mathbf{v}(T),\mathbf{w}(T)\big)_\Omega + \big(M\mathbf{v}(0),\mathbf{w}(0)\big)_\Omega - \big(\underline{A}_{\mathbf{n}}\mathbf{v},\mathbf{w}\big)_{(0,T)\times\partial\Omega}\,, \qquad \mathbf{v},\mathbf{w}\in\mathrm{C}^1(\overline{Q};\mathbb{R}^m)\,.$$

In order to define weak solutions, we include initial values for $t = 0$ and boundary conditions on Γ_k for $k = 1,\dots,m$ in the right-hand side. Therefore, we use a test space $\mathcal{V}^* \subset \mathrm{C}^1(\overline{Q};\mathbb{R}^m)$ such that

$$\big(\mathbf{v},L^*\mathbf{w}\big)_Q = \big(L\mathbf{v},\mathbf{w}\big)_Q + \big(M\mathbf{v}(0),\mathbf{w}(0)\big)_\Omega - \big(\underline{A}_{\mathbf{n}}\mathbf{v},\mathbf{w}\big)_{(0,T)\times\partial\Omega}\,, \qquad \mathbf{v}\in\mathrm{C}^1(\overline{Q};\mathbb{R}^m)\,,\ \mathbf{w}\in\mathcal{V}^*$$

with

$$\big(\underline{A}_{\mathbf{n}}\mathbf{v},\mathbf{w}\big)_{(0,T)\times\partial\Omega} = \sum_{k=1}^m \big((\underline{A}_{\mathbf{n}}\mathbf{v})_k,w_k\big)_{(0,T)\times\Gamma_k}\,, \qquad \mathbf{v}\in\mathrm{C}^1(\overline{Q};\mathbb{R}^m)\,,\ \mathbf{w} = (w_1,\dots,w_m)\in\mathcal{V}^*\,. \tag{1}$$

The property (1) characterizes adjoint boundaries $\Gamma_k^* \subset \partial\Omega$ for $k = 1,\dots,m$, so that the test space is defined by

$$\mathcal{V}^* = \big\{\mathbf{w}\in\mathrm{C}^1(\overline{Q};\mathbb{R}^m)\colon \mathbf{w}(T) = \mathbf{0} \text{ in } \Omega\,,\ \mathbf{w}(t)\in\mathcal{S}^* \text{ for } t\in[0,T)\big\}$$
$$\text{with}\quad \mathcal{S}^* = \big\{\mathbf{w}\in\mathrm{C}^1(\overline{\Omega};\mathbb{R}^m)\colon (\underline{A}_{\mathbf{n}}\mathbf{w})_k = 0 \text{ on } \Gamma_k^*\,,\ k = 1,\dots,m\big\}$$

with homogeneous final values at $t = T$ and homogenous values at the adjoint boundaries.

Our aim is to find a *weak solution* $\mathbf{u}\in\mathrm{L}_2(Q;\mathbb{R}^m)$ solving

$$\big(\mathbf{u},L^*\mathbf{w}\big)_Q = \langle\ell,\mathbf{w}\rangle \quad \text{with} \quad \langle\ell,\mathbf{w}\rangle = \big(\mathbf{f},\mathbf{w}\big)_Q + \big(M\mathbf{u}_0,\mathbf{w}(0)\big)_\Omega - \big(\mathbf{g},\mathbf{w}\big)_{(0,T)\times\partial\Omega}\,, \qquad \mathbf{w}\in\mathcal{V}^* \tag{2}$$

for given volume data $\mathbf{f}\in\mathrm{L}_2(Q;\mathbb{R}^m)$, initial data $\mathbf{u}_0\in\mathrm{L}_2(\Omega;\mathbb{R}^m)$, and boundary data $\mathbf{g}\in\mathrm{L}_2((0,T)\times\partial\Omega;\mathbb{R}^m)$, where the boundary data $\mathbf{g} = (g_k)_{k=1,\dots,m}$ are extended to $\partial\Omega$ by $g_k = 0$ on $\partial\Omega\setminus\Gamma_k$ for $k = 1,\dots,m$.

Testing the weak solution $\mathbf{u}\in\mathrm{L}_2(Q;\mathbb{R}^m)$ in (2) with functions in $\mathbf{v}\in\mathrm{C}^1_{\mathrm{c}}(Q;\mathbb{R}^m)$ defines the weak derivative $L\mathbf{u} = \mathbf{f}$ in $\mathrm{L}_2(Q;\mathbb{R}^m)$. If in addition $\mathbf{u}(0)\in\mathrm{L}_2(\Omega;\mathbb{R}^m)$ and $\underline{A}_{\mathbf{n}}\mathbf{u}|_{(0,T)\times\Gamma_k}\in\mathrm{L}_2((0,T)\times\Gamma_k)$ for $k = 1,\dots,m$, the weak solution is also a *strong solution* characterized by

$$L\mathbf{u} = \mathbf{f} \text{ in } \mathrm{L}_2(Q;\mathbb{R}^m)\,, \qquad \mathbf{u}(0) = \mathbf{u}_0 \text{ in } \mathrm{L}_2(\Omega;\mathbb{R}^m)\,, \qquad (\underline{A}_{\mathbf{n}}\mathbf{u})_k = g_k \text{ on } \mathrm{L}_2((0,T)\times\Gamma_k)\,,\ k = 1,\dots,m\,. \tag{3}$$

This is now specified for acoustic, elastic and electro-magnetic waves.

Acoustic waves. The second-order wave equation

$$\varrho\partial_t^2\phi - \nabla\cdot(\kappa\nabla\phi) = b$$

is considered as first-order system with $p = \partial_t\phi$ and $\mathbf{q} = -\kappa\nabla\phi$, i.e.,

$$\begin{aligned} \varrho\partial_t p + \nabla\cdot\mathbf{q} = b \quad &\text{and} \quad \partial_t\mathbf{q} + \kappa\nabla p = \mathbf{0} && \text{in } (0,T)\times\Omega\,,\\ p(0) = p_0 \quad &\text{and} \quad \mathbf{q}(0) = \mathbf{q}_0 && \text{in } \Omega \text{ at } t = 0\,,\\ p(t) = p_{\mathrm{D}}(t) \text{ on } \Gamma_{\mathrm{D}} \quad &\text{and} \quad \mathbf{n}\cdot\mathbf{q}(t) = g_{\mathrm{N}}(t) \text{ on } \Gamma_{\mathrm{N}} && \text{on } \partial\Omega \text{ for } t\in(0,T) \end{aligned}$$

for volume data b, boundary data g_{N}, p_{D}, initial data $\mathbf{q}_0$, p_0, positive parameters ϱ,κ, and the disjoint decomposition of the boundary $\partial\Omega = \Gamma_{\mathrm{D}}\cup\Gamma_{\mathrm{N}}$ into Dirichlet and Neumann part. The corresponding Friedrichs system with $m = 1+d$ components is given by

$$\mathbf{u} = \begin{pmatrix} p\\ \mathbf{q}\end{pmatrix}, \quad M\mathbf{u} = \begin{pmatrix} \varrho p\\ \kappa^{-1}\mathbf{q}\end{pmatrix}, \quad A\mathbf{u} = \begin{pmatrix} \nabla\cdot\mathbf{q}\\ \nabla p\end{pmatrix}, \quad \underline{A}_{\mathbf{n}}\mathbf{u} = \begin{pmatrix} \mathbf{n}\cdot\mathbf{q}\\ p\mathbf{n}\end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} b\\ \mathbf{0}\end{pmatrix}, \quad \mathbf{g} = \begin{pmatrix} g_{\mathrm{N}}\\ p_{\mathrm{D}}\mathbf{n}\end{pmatrix}, \tag{5}$$

so that for smooth functions $\varphi,\boldsymbol{\psi}$ with $\varphi = 0$ on $(0,T)\times\Gamma_{\mathrm{D}}$ and $\mathbf{n}\cdot\boldsymbol{\psi} = 0$ on $(0,T)\times\Gamma_{\mathrm{N}}$

$$\big(\underline{A}_{\mathbf{n}}(p,\mathbf{q}),(\varphi,\boldsymbol{\psi})\big)_{(0,T)\times\partial\Omega} = \big(\mathbf{n}\cdot\mathbf{q},\varphi\big)_{(0,T)\times\Gamma_{\mathrm{N}}} + \big(p,\mathbf{n}\cdot\boldsymbol{\psi}\big)_{(0,T)\times\Gamma_{\mathrm{D}}}\,.$$

In two space dimensions, this corresponds to the boundary parts $\Gamma_1 = \Gamma_1^* = \Gamma_{\mathrm{D}}$ and $\Gamma_2 = \Gamma_2^* = \Gamma_3 = \Gamma_3^* = \Gamma_{\mathrm{N}}$, and

$$\underline{M} = \begin{pmatrix} \varrho & 0 & 0\\ 0 & \kappa^{-1} & 0\\ 0 & 0 & \kappa^{-1}\end{pmatrix} \in \mathrm{L}_\infty(\Omega;\mathbb{R}^{3\times3}_{\mathrm{sym}})\,, \quad \underline{A}_1 = \begin{pmatrix} 0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 0\end{pmatrix} \in \mathbb{R}^{3\times3}_{\mathrm{sym}}\,, \quad \underline{A}_2 = \begin{pmatrix} 0 & 0 & 1\\ 0 & 0 & 0\\ 1 & 0 & 0\end{pmatrix} \in \mathbb{R}^{3\times3}_{\mathrm{sym}}\,.$$

Elastic waves. Linear elastic waves are described by the first-order system for velocity $\mathbf{v}$ and stress $\boldsymbol{\sigma}$

$$\begin{aligned} \varrho\partial_t\mathbf{v} - \nabla\cdot\boldsymbol{\sigma} = \mathbf{b} \quad &\text{and} \quad \partial_t\boldsymbol{\sigma} - \mathbf{C}\boldsymbol{\varepsilon}(\mathbf{v}) = \mathbf{0} && \text{in } (0,T)\times\Omega\,,\\ \mathbf{v}(0) = \mathbf{v}_0 \quad &\text{and} \quad \boldsymbol{\sigma}(0) = \boldsymbol{\sigma}_0 && \text{in } \Omega \text{ at } t = 0\,,\\ \mathbf{v}(t) = \mathbf{v}_{\mathrm{D}}(t) \text{ on } \Gamma_{\mathrm{D}} \quad &\text{and} \quad \boldsymbol{\sigma}\mathbf{n} = \mathbf{g}_{\mathrm{N}}(t) \text{ on } \Gamma_{\mathrm{N}} && \text{on } \partial\Omega \text{ for } t\in(0,T) \end{aligned}$$

with mass density ϱ and, in isotropic media, with $\mathbf{C}\boldsymbol{\varepsilon} = 2\mu\boldsymbol{\varepsilon} + \lambda\,\mathrm{trace}(\boldsymbol{\varepsilon})\mathbf{I}_3$ depending on the Lamé parameters $\mu,\lambda > 0$. This corresponds to the Friedrichs system with

$$\mathbf{u} = \begin{pmatrix} \mathbf{v}\\ \boldsymbol{\sigma}\end{pmatrix}, \; M\mathbf{u} = \begin{pmatrix} \varrho\mathbf{v}\\ \mathbf{C}^{-1}\boldsymbol{\sigma}\end{pmatrix}, \; A\mathbf{u} = \begin{pmatrix} -\nabla\cdot\boldsymbol{\sigma}\\ -\boldsymbol{\varepsilon}(\mathbf{v})\end{pmatrix}, \; \underline{A}_{\mathbf{n}}\mathbf{u} = \begin{pmatrix} -\boldsymbol{\sigma}\mathbf{n}\\ -\mathbf{n}\mathbf{v}^\top - \mathbf{v}\mathbf{n}^\top\end{pmatrix}, \; \mathbf{f} = \begin{pmatrix} \mathbf{b}\\ \mathbf{0}\end{pmatrix}, \; \mathbf{g} = \begin{pmatrix} -\mathbf{g}_{\mathrm{N}}\\ -\mathbf{n}\mathbf{v}_{\mathrm{D}}^\top - \mathbf{v}_{\mathrm{D}}\mathbf{n}^\top\end{pmatrix}. \tag{6}$$

For $d = 3$ we have $m = 9$ and $\Gamma_k = \Gamma_k^* = \Gamma_{\mathrm{D}}$ for $k = 1,2,3$, and $\Gamma_k = \Gamma_k^* = \Gamma_{\mathrm{N}}$ for $k = 4,\dots,9$.

Electro-magnetic waves. The first-order system for the electric field $\mathbf{E}$ and the magnetic field intensity $\mathbf{H}$

$$\begin{aligned} \varepsilon\partial_t\mathbf{E} - \nabla\times\mathbf{H} = -\mathbf{J} \quad &\text{and} \quad \mu\partial_t\mathbf{H} + \nabla\times\mathbf{E} = \mathbf{0} && \text{in } (0,T)\times\Omega\,,\\ \mathbf{E}(0) = \mathbf{E}_0 \quad &\text{and} \quad \mathbf{H}(0) = \mathbf{H}_0 && \text{in } \Omega \text{ at } t = 0\,,\\ \mathbf{n}\times\mathbf{E}(t) = \mathbf{0} \text{ on } \Gamma_{\mathrm{E}} \quad &\text{and} \quad \mathbf{n}\times\mathbf{H}(t) = \mathbf{g}_{\mathrm{M}} \text{ on } \Gamma_{\mathrm{M}} && \text{on } \partial\Omega \text{ for } t\in(0,T) \end{aligned}$$

with permittivity ε, permeability μ, and boundary decomposition $\partial\Omega = \Gamma_{\mathrm{E}}\cup\Gamma_{\mathrm{M}}$ corresponds to a Friedrichs system with

$$\mathbf{u} = \begin{pmatrix} \mathbf{E}\\ \mathbf{H}\end{pmatrix}, \quad M\mathbf{u} = \begin{pmatrix} \varepsilon\mathbf{E}\\ \mu\mathbf{H}\end{pmatrix}, \quad A\mathbf{u} = \begin{pmatrix} -\nabla\times\mathbf{H}\\ \nabla\times\mathbf{E}\end{pmatrix}, \quad \underline{A}_{\mathbf{n}}\mathbf{u} = \begin{pmatrix} -\mathbf{n}\times\mathbf{H}\\ \mathbf{n}\times\mathbf{E}\end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} -\mathbf{J}\\ \mathbf{0}\end{pmatrix}, \quad \mathbf{g} = \begin{pmatrix} -\mathbf{g}_{\mathrm{M}}\\ \mathbf{0}\end{pmatrix}. \tag{7}$$

For $d = 3$ we have $m = 6$ and $\Gamma_k = \Gamma_k^* = \Gamma_{\mathrm{E}}$ for $k = 1,2,3$, and $\Gamma_k = \Gamma_k^* = \Gamma_{\mathrm{M}}$ for $k = 4,5,6$.

Remark 1. *We only consider the case that the symmetric matrices* $\underline{A}_j$*,* $j = 1, \dots, d$*, are constant in* Ω*. In general,* $\underline{A}_j$ *may depend on* $\mathbf{x} \in \Omega$*, e.g., for the linear transport equation* $Lu = \partial_t u + \mathbf{a} \cdot \nabla u$ *with* $m = 1$ *and transport vector* $\mathbf{a}(\mathbf{x}) \in \mathbb{R}^d$*. Then,* Γ_1 *is the inflow boundary, and for the adjoint equation we obtain* $L^* v = -\partial_t v - \mathbf{a} \cdot \nabla v - (\nabla \cdot \mathbf{a}) v$ *with* $\Gamma_1^* = \partial\Omega \setminus \Gamma_1$*. For the DG analysis of this case we refer to [Di Pietro and Ern, 2011, Chap. 2] in the steady case and to [Dörfler et al., 2016] for a Petrov–Galerkin space-time method.*

The suitable choice of the subsets $\Gamma_k \subset \partial\Omega$ for $k = 1, \dots, m$ for the boundary conditions in general Friedrichs systems is discussed in [Di Pietro and Ern, 2011, Chap. 7.2]. Here we consider the special case for wave systems. The property (1) characterizes the adjoint boundaries $\Gamma_k^* \subset \partial\Omega$ for $k = 1, \dots, m$, and we observe

$$\sum_{k=1}^m \big((\underline{A}_\mathbf{n}\mathbf{v})_k, w_k\big)_{(0,T)\times\Gamma_k} = \big(\underline{A}_\mathbf{n}\mathbf{v}, \mathbf{w}\big)_{(0,T)\times\partial\Omega} = \big(\mathbf{v}, \underline{A}_\mathbf{n}\mathbf{w}\big)_{(0,T)\times\partial\Omega} = \sum_{k=1}^m \big(v_k, (\underline{A}_\mathbf{n}\mathbf{w})_k\big)_{(0,T)\times\partial\Omega\setminus\Gamma_k^*}$$

for $\mathbf{v} = (v_1 \dots, v_m) \in \mathrm{C}^1(Q; \mathbb{R}^m)$ and $\mathbf{w} = (w_1, \dots, w_m) \in \mathcal{V}^*$ and thus, defining

$$\mathcal{V} = \big\{\mathbf{v} \in \mathrm{C}^1(\overline{Q}; \mathbb{R}^m) \colon \mathbf{v}(0) = \mathbf{0} \text{ in } \Omega\,,\ (\underline{A}_\mathbf{n}\mathbf{v})_k = 0 \text{ on } (0,T] \times \Gamma_k\,,\ k = 1, \dots, m\big\}$$

with homogeneous initial value at $t = 0$ and homogeneous boundary values on Γ_k, we obtain

$$\big(\underline{A}_\mathbf{n}\mathbf{v}, \mathbf{w}\big)_{(0,T)\times\partial\Omega} = \big(\mathbf{v}, \underline{A}_\mathbf{n}\mathbf{w}\big)_{(0,T)\times\partial\Omega} = 0\,, \qquad \mathbf{v} \in \mathcal{V}\,,\ \mathbf{w} \in \mathcal{V}^*\,.$$

Boundary conditions are required in order to obtain uniqueness and well-posedness of the solution. Therefore, we require for the subsets $\Gamma_k \subset \partial\Omega$, for $k = 1, \dots, m$, that the operators L and L^* are injective on $\mathcal{V}$ and $\mathcal{V}^*$, respectively, i.e.,

$$\big\{\mathbf{v} \in \mathcal{V} \colon L\mathbf{v} = \mathbf{0}\big\} = \{\mathbf{0}\}\,, \qquad \big\{\mathbf{w} \in \mathcal{V}^* \colon L^*\mathbf{w} = \mathbf{0}\big\} = \{\mathbf{0}\}\,, \tag{8}$$

where the relatively open adjoint boundaries $\Gamma_k^* \subset \partial\Omega$ for $k = 1, \dots, m$ are determined by property (1).

Now we show that both conditions in (8) are necessary. The first condition for Γ_k is required for uniqueness for strong solutions: if $\mathbf{v} \in \mathcal{V} \setminus \{\mathbf{0}\}$ exists with $L\mathbf{v} = \mathbf{0}$, then this is a non-trivial homogeneous strong solution, i.e., $\mathbf{v}$ solves (3) with $\mathbf{u}_0 = \mathbf{0}$, $\mathbf{f} = \mathbf{0}$, and $\mathbf{g} = \mathbf{0}$. On the other hand, if the second condition is violated, weak solutions do not exist for all volume data: if $\mathbf{w} \in \mathcal{V}^* \setminus \{\mathbf{0}\}$ and $\mathbf{f} \in \mathrm{L}_2(Q; \mathbb{R}^m)$ exists with $L^*\mathbf{w} = \mathbf{0}$ and $(\mathbf{f}, \mathbf{w})_Q \neq 0$, no weak solution of (2) with homogeneous initial and boundary data $\mathbf{u}_0 = \mathbf{0}$ and $\mathbf{g} = \mathbf{0}$ exists.

Remark 2. *The formulation of wave equations in our examples as Friedrichs systems yields symmetric matrices of the form* $\underline{A}_j = \begin{pmatrix} \underline{0} & \underline{\tilde{A}}_j \\ \underline{\tilde{A}}_j^\top & \underline{0} \end{pmatrix}$ *with* $\underline{\tilde{A}}_j \in \mathbb{R}^{m_1 \times m_2}$ *and* $m = m_1 + m_2$*. For the boundary conditions we can select a relatively open set* $\Gamma_1 \subset \partial\Omega$*. Then, defining* $\Gamma_k = \Gamma_1$ *for* $k = 2, \dots, m_1$*,* $\Gamma_k = \partial\Omega \setminus \overline{\Gamma}_1$ *for* $k = m_1 + 1, \dots, m$*, and* $\Gamma_k^* = \Gamma_k$ *for* $k = 1, \dots, m$*, we observe that property (1) and conditions (8) are satisfied.*

Remark 3. *For smooth domains and data, the solution is also smooth, e.g., for acoustics* $\phi(t) \in \mathrm{H}^s(\Omega)$ *for all* $t \in [0,T]$ *with* $s \geq 2$*. This allows for improved approximation orders* $\mathcal{O}(h^s)$ *for* ϕ*. On the other hand, the necessary regularity requirements are quite restrictive [Rauch, 1985], and the second-order formulation does not allow for the convergence analysis of piece-wise discontinuous solutions.*

Remark 4. *Waves in real media are dissipative and dispersive; e.g., modeling electro-magnetic waves in matter needs to include conductivity and impedance. The DG analysis can be extended to this case; see, e.g., [Di Pietro and Ern, 2011, Chap. 7] for the steady case and [Dörfler et al., 2020] for visco-elastic waves with impedance boundary conditions. In the elastic model for Rayleigh damping or for the Kelvin–Voigt model, the linear operator takes the form* $L = M\partial_t + D + A$ *with* $(D\mathbf{v})(t, \mathbf{x}) = \underline{D}(\mathbf{x})\mathbf{v}(t, \mathbf{x})$ *and* $\underline{D} \in \mathrm{L}_\infty(\Omega; \mathbb{R}^{d\times d}_{sym})$ *symmetric positive semi-definite; then,* $L^* = -M\partial_t + D - A$*. All our subsequent results extend to this case, but for simplicity we only consider the case* $\underline{D} = \underline{0}$*.*

3. The full-upwind discontinuous Galerkin discretization

In this section we introduce an upwind DG discretization for the first-order system.

3.1. The DG finite element space in the space-time cylinder

For the discretization, we use tensor product space-time cells combining the mesh in space with a decomposition in time. For $0 = t_0 < t_1 < \cdots < t_N = T$, we define time intervals $I_{n,h} = (t_{n-1}, t_n)$, time-step sizes $\triangle t_n = t_n - t_{n-1}$, and

$$I_h = (t_0, t_1) \cup \cdots \cup (t_{N-1}, t_N) \subset I = (0, T) , \qquad \partial I_h = \{t_0, t_1, \ldots, t_{N-1}, t_N\} .$$

We set $\triangle t = \max \triangle t_n$, and we assume quasi-uniformity, i.e., $\triangle t_n \in [C_{\rm sr} \triangle t, \triangle t]$ with $C_{\rm sr} \in (0, 1]$ independent of N.

Let $\mathcal{K}_h$ be a mesh so that $\Omega_h = \bigcup_{K \in \mathcal{K}_h} K$ is a decomposition in space into open cells $K \subset \Omega \subset \mathbb{R}^d$. Then, we obtain a tensor-product decomposition into space-time cells $R = I_{n,h} \times K$

$$Q_h = I_h \times \Omega_h = \bigcup_{n=1}^{N} Q_{n,h} = \bigcup_{R \in \mathcal{R}_h} R \ \subset \ Q = I \times \Omega \subset \mathbb{R}^{1+d} , \qquad Q_{n,h} = \bigcup_{K \in \mathcal{K}_h} I_{n,h} \times K \ \subset \ I_{n,h} \times \Omega$$

of the space-time cylinder Q. Let $F \in \mathcal{F}_K$ be the faces of the element K, and we set $\mathcal{F}_h = \bigcup_K \mathcal{F}_K$, so that $\partial \Omega_h = \overline{\bigcup_{F \in \mathcal{F}_h} F}$ is the skeleton in space; $\partial Q_h = \bigcup_{n=0}^{N} \{t_n\} \times \partial \Omega_h$ is the corresponding space-time skeleton. For inner faces $F \in \mathcal{F}_h \cap \Omega$ and $K \in \mathcal{K}_h$, let K_F be the neighboring cell such that $\overline{F} = \partial K \cap \partial K_F$. On boundary faces $F \in \mathcal{F}_h \cap \partial \Omega$ we set $K_F = K$. Let $\mathbf{n}_K$ be the outer unit normal vector on ∂K. We assume that $\overline{\Omega} = \Omega_h \cup \partial \Omega_h$ and that the boundary decomposition is compatible with the mesh, i.e., $\overline{\Gamma}_k = \overline{\bigcup_{F \in \mathcal{F}_K \cap \Gamma_k} F}$ for $k = 1, \ldots, m$.

We set $h_K = \operatorname{diam} K$, $h_F = \operatorname{diam} F$, and $h = \max h_K$. We assume quasi-uniform meshes and shape-regularity, i.e., $h_F \geq C_{\rm sr} h_K$ for $F \in \mathcal{F}_K$ with $C_{\rm sr} > 0$ independent of h_K. In the following, we use the mesh-dependent norms

$$\big\| h^{\alpha/2} \mathbf{v}_h \big\|_Q = \Big(\sum_{n=1}^{N} \sum_{K \in \mathcal{K}_h} h_K^{\alpha} \|\mathbf{v}_h\|^2_{I_{n,h} \times K} \Big)^{1/2} , \qquad \alpha \in \mathbb{R} . \tag{9}$$

In order to calibrate the accuracy in space and time, we assume, depending on a reference velocity $c_{\rm ref} > 0$, that the mesh size in time and space are well balanced satisfying

$$c_{\rm ref} \triangle t \leq h . \tag{10}$$

Remark 5. *For simplicity we use only tensor-product space-time meshes. For the extension to more general meshes in the space-time cylinder we refer to [Gopalakrishnan et al., 2017], see also the analysis in [Bansal et al., 2021]. General meshes in $\mathbb{R}^{1+d}$ are considered in [Schafelner, 2022]. Then, the condition (10) can be relaxed to a local condition.*

The DG discretization is defined for a finite dimensional subspace $V_h \subset \mathcal{V}_h \subset \mathrm{C}^1(I_h; \mathcal{S}_h)$, where

$$\begin{aligned} \mathcal{V}_h &= \big\{ \mathbf{v}_h \in \mathrm{C}^1(Q_h; \mathbb{R}^m) \colon \mathbf{v}_{n,h,K} = \mathbf{v}_h|_{I_{n,h} \times K} \text{ extends continuously to } \mathbf{v}_{n,h,K} \in \mathrm{C}^0(\overline{I_{n,h} \times K}; \mathbb{R}^m) \big\} , \\ \mathcal{S}_h &= \big\{ \mathbf{v}_h \in \mathrm{C}^1(\Omega_h; \mathbb{R}^m) \colon \mathbf{v}_{h,K} = \mathbf{v}_h|_K \text{ extends continuously to } \mathbf{v}_{h,K} \in \mathrm{C}^0(\overline{K}; \mathbb{R}^m) \big\} . \end{aligned}$$

For the positive definite matrix function $\underline{M} \in \mathrm{L}_\infty(\Omega; \mathbb{R}^{m \times m}_{\rm sym})$ let $\underline{M}_h \in \mathrm{L}_\infty(\Omega_h; \mathbb{R}^{m \times m}_{\rm sym})$ be a piecewise constant approximation, and for $K \in \mathcal{K}_h$ let $\underline{M}_{h,K} \in \mathbb{R}^{m \times m}_{\rm sym}$ be the continuous extension of $\underline{M}_h|_K$ to $\overline{K}$; in case of material jumps this can result to different values on the left and right side of a face, i.e., $M_K|_F \neq M_{K_F}|_F$.

Let $L_h = M_h \partial_t + A$ be the corresponding linear differential operator, where the approximated operator M_h is given by $(M_h \mathbf{v})(t, \mathbf{x}) = \underline{M}_h(\mathbf{x}) \mathbf{v}(t, \mathbf{x})$. Note that then $L_h(V_h) \subset V_h$.

For our applications, we use a tensor-product construction of the finite element space.

For every space-time cell $R = I_{n,h} \times K$ we select polynomial degrees $p_R = p_{n,K} \geq 0$ in time and $q_R = q_{n,K} \geq 0$ in space. With this we define the discontinuous finite element spaces

$$S_{n,h} = \prod_{K \in \mathcal{K}_h} \mathbb{P}_{q_{n,K}}(K; \mathbb{R}^m) \subset \mathcal{S}_h \,, \qquad S_h = S_{1,h} + \cdots + S_{N,h} \subset \mathcal{S}_h \,, \tag{11a}$$

$$V_{n,h} = \prod_{K \in \mathcal{K}_h} \mathbb{P}_{p_{n,K}} \otimes \mathbb{P}_{q_{n,K}}(K; \mathbb{R}^m) \subset \mathcal{V}_{n,h} \,, \qquad V_h = V_{1,h} + \cdots + V_{N,h} \subset \mathcal{V}_h \,, \tag{11b}$$

where $\mathbb{P}_p$ denotes the set of polynomials up to order p. For the following, we fix $p = \max p_R$ and $q = \max q_R$, so that

$$S_{n,h} \subset S_h \subset \mathbb{P}_q(\Omega_h; \mathbb{R}^m) \subset \mathcal{S}_h \,, \qquad V_h \subset \mathbb{P}_p(I_h) \otimes S_h \subset \mathbb{P}_p(I_h) \otimes \mathbb{P}_q(\Omega_h; \mathbb{R}^m) \subset \mathcal{V}_h \,.$$

On the space-time skeleton $\partial Q_h = \bigcup_{n=0}^N \{t_n\} \times \overline{\Omega} \,\cup\, I_h \times \partial\Omega_h$, the inverse inequality and the discrete trace inequality [Di Pietro and Ern, 2011, Lem. 1.44 and Lem. 1.46] yield

$$\big\| h^{1/2} M_h^{-1/2} L_h \mathbf{v}_h \big\|_{Q_h} \leq C_{\mathrm{inv}} \big\| h^{-1/2} M_h^{1/2} \mathbf{v}_h \big\|_Q \,, \tag{12a}$$

$$\big\| M_h^{1/2} \mathbf{v}_h \big\|_{\partial Q_h} \leq C_{\mathrm{tr}} \big\| h^{-1/2} M_h^{1/2} \mathbf{v}_h \big\|_Q \,, \qquad \mathbf{v}_h \in V_h \,, \tag{12b}$$

with $C_{\mathrm{inv}}, C_{\mathrm{tr}} > 0$ depending on the space-time mesh regularity (and thus also on c_{ref}), the polynomial degrees in V_h, and the material parameters.

Let $\Pi_h \colon \mathrm{L}_2(Q; \mathbb{R}^m) \longrightarrow V_h$ be the space-time L_2 projection defined by

$$\big(M_h \Pi_h \mathbf{v}, \mathbf{v}_h \big)_Q = \big(M_h \mathbf{v}, \mathbf{v}_h \big)_Q \,, \qquad \mathbf{v}_h \in V_h \,. \tag{13}$$

For $\mathbf{v}_h \in V_h$, let $\mathbf{v}_{n,h} \in \mathrm{C}^0\big([t_{n-1}, t_n]; \mathrm{L}_2(\Omega, h; \mathbb{R}^m)\big)$ be the extension of $\mathbf{v}_h|_{Q_{n,h}} \in \mathrm{L}_2(Q_{n,h}; \mathbb{R}^m)$ to $[t_{n-1}, t_n]$.

In every time interval $I_{n,h}$ we use the projection $\Pi_{n,h} \colon \mathrm{L}_2(\Omega; \mathbb{R}^m) \longrightarrow S_{n,h} \subset S_h$ defined by

$$\big(M_h \Pi_{n,h} \mathbf{w}, \mathbf{w}_{n,h} \big)_\Omega = \big(M_h \mathbf{w}_n, \mathbf{w}_{n,h} \big)_\Omega \,, \qquad \mathbf{w}_{n,h} \in S_{n,h} \,.$$

3.2. A discontinuous Galerkin method in time

For $\mathbf{v}_h, \mathbf{w}_h \in \mathcal{V}_h$ we obtain after integration by parts in all intervals $I_{n,h} \subset I_h$

$$\big(M_h \partial_t \mathbf{v}_h, \mathbf{w}_h\big)_{Q_h} = \sum_{n=1}^{N} \Big(-\big(M_h \mathbf{v}_{n,h}, \partial_t \mathbf{w}_{n,h}\big)_{Q_{n,h}} + \big(M_h \mathbf{v}_{n,h}(t_n), \mathbf{w}_{n,h}(t_n)\big)_\Omega - \big(M_h \mathbf{v}_{n,h}(t_{n-1}), \mathbf{w}_{n,h}(t_{n-1})\big)_\Omega \Big) .$$

Introducing the jump terms $[\mathbf{w}_h]_n = \mathbf{w}_{n+1,h}(t_n) - \mathbf{w}_{n,h}(t_n)$ for $n = 1, \dots, N-1$ and $[\mathbf{w}_h]_N = -\mathbf{w}_{N,h}(t_N)$, we define the dual representation of the full upwind DG method in time

$$m_h(\mathbf{v}_h, \mathbf{w}_h) = -\big(M_h \mathbf{v}_{n,h}, \partial_t \mathbf{w}_{n,h}\big)_{Q_h} - \sum_{n=1}^{N} \big(M_h \mathbf{v}_{n,h}(t_n), [\mathbf{w}_h]_n\big)_\Omega , \qquad \mathbf{v}_h, \mathbf{w}_h \in \mathcal{V}_h . \tag{14}$$

We have dual consistency by construction, i.e.,

$$m_h(\mathbf{v}_h, \mathbf{w}) = -\big(M_h \mathbf{v}_h, \partial_t \mathbf{w}\big)_{Q_h} , \qquad \mathbf{w} \in \mathcal{V}^* . \tag{15}$$

Again integrating by parts and defining $[\mathbf{v}_h]_0 = \mathbf{v}_{1,h}(0)$ yields the primal representation

$$m_h(\mathbf{v}_h, \mathbf{w}_h) = \big(M_h \partial_t \mathbf{v}_h, \mathbf{w}_h\big)_{Q_h} + \sum_{n=1}^{N} \big(M_h [\mathbf{v}_h]_{n-1}, \mathbf{w}_{n,h}(t_{n-1})\big)_\Omega . \tag{16}$$

Together, we obtain

$$\begin{aligned} 2\, m_h(\mathbf{v}_h, \mathbf{v}_h) &= m_h(\mathbf{v}_h, \mathbf{v}_h) + m_h(\mathbf{v}_h, \mathbf{v}_h) \\ &= \sum_{n=1}^{N} \Big(\big(M_h [\mathbf{v}_h]_{n-1}, \mathbf{v}_{n,h}(t_{n-1})\big)_\Omega - \big(M_h \mathbf{v}_{n,h}(t_n), [\mathbf{v}_h]_n\big)_\Omega \Big) \\ &= \big(M_h \mathbf{v}_h(0), \mathbf{v}_h(0)\big)_\Omega + \sum_{n=1}^{N-1} \Big(\big(M_h [\mathbf{v}_h]_n, \mathbf{v}_{n+1,h}(t_n)\big)_\Omega - \big(M_h \mathbf{v}_{n,h}(t_n), [\mathbf{v}_h]_n\big)_\Omega \Big) + \big(M_h \mathbf{v}_h(T), \mathbf{v}_h(T)\big)_\Omega , \end{aligned}$$

which yields

$$m_h(\mathbf{v}_h, \mathbf{v}_h) = \frac{1}{2} \sum_{n=0}^{N} \big(M_h [\mathbf{v}_h]_n, [\mathbf{v}_h]_n\big)_\Omega \ge 0 , \qquad \mathbf{v}_h \in \mathcal{V}_h , \tag{17}$$

so that

$$m_h\big(\mathbf{v}_h, \mathbf{v}_h\big) = 0 \quad \Longrightarrow \quad m_h\big(\mathbf{v}_h, \mathbf{w}_h\big) = -\big(M_h \mathbf{v}_h, \partial_t \mathbf{w}\big)_{Q_h} = \big(M_h \partial_t \mathbf{v}_h, \mathbf{w}\big)_{Q_h} , \qquad \mathbf{v}_h, \mathbf{w}_h \in \mathcal{V}_h . \tag{18}$$

For $m_h(\mathbf{v}_h,\mathbf{v}_h)=0$ we observe $\mathbf{v}_h\in \mathrm{H}^1_0(0,T;\mathcal{S}_h)$. This yields with $d_T(t)=T-t$

$$\begin{aligned}\big(M_h\mathbf{v}_h,\mathbf{v}_h\big)_Q &= \int_0^T \big(M_h\mathbf{v}_h(t),\mathbf{v}_h(t)\big)_\Omega\,\mathrm{d}t = -\int_0^T \big(M_h\mathbf{v}_h(t),\mathbf{v}_h(t)\big)_\Omega \partial_t d_T(t)\,\mathrm{d}t\\ &= 2\int_0^T \big(M_h\partial_t\mathbf{v}_h(t),\mathbf{v}_h(t)\big)_\Omega d_T(t)\,\mathrm{d}t \le 2T\,\big\|M_h^{-1/2}\partial_t\mathbf{v}_h\big\|_{Q_h}\big\|M_h^{1/2}\mathbf{v}_h\big\|_Q\,,\end{aligned}$$

i.e., we have $\big\|M_h^{1/2}\mathbf{v}_h\big\|_Q\le 2T\,\big\|M_h^{-1/2}\partial_t\mathbf{v}_h\big\|_{Q_h}$. This extends to discontinuous functions in $\mathcal{V}_h$ as follows.

Lemma 6. *We have*

$$\big(M_h\mathbf{v}_h,\mathbf{v}_h\big)_Q + \sum_{n=0}^{N-1} d_T(t_n)\big(M_h[\mathbf{v}_h]_n,[\mathbf{v}_h]_n\big)_\Omega \le 2\,m_h(\mathbf{v}_h,d_T\mathbf{v}_h)\,,\qquad \mathbf{v}_h\in\mathcal{V}_h\,.$$

Proof. The assertion follows from

$$\begin{aligned}\big(M_h\mathbf{v}_h,\mathbf{v}_h\big)_Q &= -\sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} \big(M_h\mathbf{v}_h(t),\mathbf{v}_h(t)\big)_\Omega\partial_t d_T(t)\,\mathrm{d}t\\ &= 2\int_0^T \big(M_h\partial_t\mathbf{v}_h(t),\mathbf{v}_h(t)\big)_\Omega d_T(t)\,\mathrm{d}t\\ &\qquad -\sum_{n=1}^{N}\Big(d_T(t_n)\big(M_h\mathbf{v}_{n,h}(t_n),\mathbf{v}_{n,h}(t_n)\big)_\Omega - d_T(t_{n-1})\big(M_h\mathbf{v}_{n,h}(t_{n-1}),\mathbf{v}_{n,h}(t_{n-1})\big)_\Omega\Big)\\ &= 2\big(M_h\partial_t\mathbf{v}_h,d_T\mathbf{v}_h\big)_{Q_h}\\ &\qquad +\sum_{n=1}^{N-1} d_T(t_n)\Big(\big(M_h\mathbf{v}_{n+1,h}(t_n),\mathbf{v}_{n+1,h}(t_n)\big)_\Omega - \big(M_h\mathbf{v}_{n,h}(t_n),\mathbf{v}_{n,h}(t_n)\big)_\Omega\Big) - T\,\|M_h^{1/2}\mathbf{v}_{1,h}(0)\|_\Omega^2\\ &\le 2\big(M_h\partial_t\mathbf{v}_h,d_T\mathbf{v}_h\big)_{Q_h} + 2\sum_{n=1}^{N-1} d_T(t_n)\big(M_h[\mathbf{v}_h]_n,\mathbf{v}_{n+1,h}(t_n)\big)_\Omega\\ &\qquad -\sum_{n=1}^{N-1} d_T(t_n)\big(M_h[\mathbf{v}_h]_n,[\mathbf{v}_h]_n\big)_\Omega - T\,\|M_h^{1/2}\mathbf{v}_{1,h}(0)\|_\Omega^2\\ &\le 2\,m_h\big(\mathbf{v}_h,d_T\mathbf{v}_h\big) - \sum_{n=0}^{N-1} d_T(t_n)\big(M_h[\mathbf{v}_h]_n,[\mathbf{v}_h]_n\big)_\Omega\end{aligned}$$

using

$$\begin{aligned}&\big(M_h\mathbf{v}_{n+1,h}(t_n),\mathbf{v}_{n+1,h}(t_n)\big)_\Omega - \big(M_h\mathbf{v}_{n,h}(t_n),\mathbf{v}_{n,h}(t_n)\big)_\Omega = \big(M_h(\mathbf{v}_{n+1,h}(t_n)-\mathbf{v}_{n,h}(t_n)),\mathbf{v}_{n+1,h}(t_n)+\mathbf{v}_{n,h}(t_n)\big)_\Omega\\ &\quad = \big(M_h[\mathbf{v}_h]_n,\mathbf{v}_{n+1,h}(t_n)\big)_\Omega + \big(M_h[\mathbf{v}_h]_n,\mathbf{v}_{n,h}(t_n)\big)_\Omega = 2\big(M_h[\mathbf{v}_h]_n,\mathbf{v}_{n+1,h}(t_n)\big)_\Omega - \big(M_h[\mathbf{v}_h]_n,[\mathbf{v}_h]_n\big)_\Omega\,.\end{aligned}$$

□

3.3. A discontinuous Galerkin method in space

For $\mathbf{v}_h, \mathbf{w}_h \in \mathcal{S}_h$ we observe, integrating by parts for all elements $K \in \mathcal{K}_h$,

$$\big(A\mathbf{v}_h, \mathbf{w}_h\big)_{\Omega_h} = \sum_{K\in\mathcal{K}_h} \Bigg(-\big(\mathbf{v}_{h,K}, A\mathbf{w}_{h,K}\big)_K + \sum_{F\in\mathcal{F}_K} \big(\underline{A}_{\mathbf{n}_K}\mathbf{v}_{h,K}, \mathbf{w}_{h,K}\big)_F \Bigg) .$$

For conforming functions $\mathbf{v}$, we have for the flux $\underline{A}_{\mathbf{n}_K}\mathbf{v} = -\underline{A}_{\mathbf{n}_{K_F}}\mathbf{v}$ on inner faces $F \subset \Omega$, and for discontinuous functions we define the jump term $[\mathbf{w}_h]_{K,F} = \mathbf{w}_{h,K_F} - \mathbf{w}_{h,K}$. On boundary faces $F \subset \partial\Omega$ this depends on the boundary conditions, and we set $(\underline{A}_\mathbf{n}[\mathbf{v}_h])_k = -2(\underline{A}_\mathbf{n}\mathbf{v}_h)_k$ on $\Gamma_k \subset \partial\Omega$ and $(\underline{A}_\mathbf{n}[\mathbf{v}_h])_k = 0$ on $\partial\Omega \setminus \Gamma_k$ for $k = 1, \dots, m$.
We use the discontinuous Galerkin with full upwind discretization in space which is of the form

$$a_h(\mathbf{v}_h, \mathbf{w}_h) = -\big(\mathbf{v}_h, A\mathbf{w}_h\big)_{\Omega_h} + \sum_{K\in\mathcal{K}_h}\sum_{F\in\mathcal{F}_K} \big(\mathbf{v}_{h,K}, \underline{A}^{\mathrm{up}}_{\mathbf{n}_K}[\mathbf{w}_h]_{K,F}\big)_F ,$$

where the upwind flux $\underline{A}^{\mathrm{up}}_{\mathbf{n}_K} \in \mathbb{R}^{m\times m}$ is obtained by solving local Riemann problems.
For the DG method we require dual consistency for the bilinear form and the right hand side for the boundary values

$$a_h\big(\mathbf{v}_h, \mathbf{w}\big) = -\big(\mathbf{v}_h, A\mathbf{w}\big)_{\Omega_h} \quad\text{and}\quad \big\langle \ell_{\partial\Omega,h}(t), \mathbf{w}\big\rangle = \big(\mathbf{g}(t), \mathbf{w}\big)_{\partial\Omega} , \qquad \mathbf{v}_h \in \mathcal{S}_h ,\ \mathbf{w} \in \mathcal{S}^* , \tag{19}$$

and for the inconsistency complement we require that $C_1 \geq c_1 > 0$ exists such that

$$c_1\big\|\underline{A}_\mathbf{n}[\mathbf{v}_h]\big\|^2_{\partial\Omega_h} \leq a_h\big(\mathbf{v}_h, \mathbf{v}_h\big) \leq C_1\big\|\underline{A}_\mathbf{n}[\mathbf{v}_h]\big\|^2_{\partial\Omega_h} , \qquad \mathbf{v}_h \in \mathcal{S}_h , \tag{20}$$

so that

$$a_h\big(\mathbf{v}_h, \mathbf{v}_h\big) = 0 \quad\Longrightarrow\quad a_h\big(\mathbf{v}_h, \mathbf{w}_h\big) = -\big(\mathbf{v}_h, A\mathbf{w}_h\big)_{\Omega_h} = \big(A\mathbf{v}_h, \mathbf{w}_h\big)_{\Omega_h} , \qquad \mathbf{v}_h, \mathbf{w}_h \in \mathcal{S}_h . \tag{21}$$

We assume that $C_1 > 0$ only depends on the material parameters, and that

$$\big|a_h\big(\mathbf{v}_h, \mathbf{w}_h\big) + \big(\mathbf{v}_h, A\mathbf{w}_h\big)_{\Omega_h}\big| \leq C_1\big\|M_h^{1/2}\mathbf{v}_h\big\|_{\partial\Omega_h}\big\|\underline{A}_\mathbf{n}[\mathbf{w}_h]\big\|_{\partial\Omega_h} , \qquad \mathbf{v}_h, \mathbf{w}_h \in \mathcal{S}_h , \tag{22a}$$

$$\big|a_h\big(\mathbf{v}_h, \mathbf{w}_h\big) + \big(A\mathbf{v}_h, \mathbf{w}_h\big)_{\Omega_h}\big| \leq C_1\big\|\underline{A}_\mathbf{n}[\mathbf{v}_h]\big\|_{\partial\Omega_h}\big\|M_h^{1/2}\mathbf{w}_h\big\|_{\partial\Omega_h} , \qquad \mathbf{v}_h, \mathbf{w}_h \in \mathcal{S}_h , \tag{22b}$$

$$\big|\big\langle \ell_{\partial\Omega,h}(t), \mathbf{w}_h\big\rangle - \big(\mathbf{g}(t), \mathbf{w}_h\big)_{\partial\Omega}\big| \leq C_1\big\|\mathbf{g}(t)\big\|_{\partial\Omega_h}\big\|M_h^{1/2}\mathbf{w}_h\big\|_{\partial\Omega_h} , \qquad \mathbf{w}_h \in \mathcal{S}_h . \tag{22c}$$

For acoustic, elastic and electro-magnetic waves the upwind flux is explicitly evaluated, e.g., in [Hochbruck et al., 2015, Sect. 4.3]. Here, we only consider the dual representation; integration by parts yields the primal representation.

Acoustic waves. The full upwind DG approximation for the acoustic wave equation (5) is given by

$$\begin{aligned} a_h\big((p_h, \mathbf{q}_h), (\varphi_h, \boldsymbol{\psi}_h)\big) = \sum_{K\in\mathcal{K}_h} \Bigg(&- \big(\mathbf{q}_{h,K}, \nabla\varphi_{h,K}\big)_K - \big(p_{h,K}, \nabla\cdot\boldsymbol{\psi}_{h,K}\big)_K \\ &- \sum_{F\in\mathcal{F}_K} \frac{1}{Z_K + Z_{K_F}} \big(p_{K,h} + Z_{K_F}\mathbf{n}_K\cdot\mathbf{q}_{K,h}, [\varphi_h]_{K,F} + Z_K\mathbf{n}_K\cdot[\boldsymbol{\psi}_h]_{K,F}\big)_F \Bigg) \end{aligned} \tag{23}$$

for $(p_h, \mathbf{q}_h), (\varphi_h, \boldsymbol{\psi}_h) \in \mathcal{S}_h$ with impedance $Z_K = \sqrt{\kappa_{h,K}\varrho_{h,K}}$ depending on the piecewise constant approximations for the material parameters $\kappa, \varrho > 0$. On inner boundaries material discontinuities can result in $Z_K \neq Z_{K_F}$, on boundary faces we define $Z_h = Z_K$ on $\partial\Omega \cap \partial K$. On Dirichlet boundary faces $F \in \mathcal{F}_h \cap \Gamma_\mathrm{D}$, we set $[p_h]_{K,F} = -2p_h$ and $\mathbf{n}\cdot[\mathbf{q}_h]_{K,F} = 0$. On Neumann boundary faces $F \in \mathcal{F}_h \cap \Gamma_\mathrm{N}$, we set $[p_h]_{K,F} = 0$ and $\mathbf{n}\cdot[\mathbf{q}_h]_{K,F} = -2\,\mathbf{n}\cdot\mathbf{q}_h$. The right-hand side is complemented by the stabilization, so that

$$\big\langle \ell_{\partial\Omega,h}(t), (\varphi_h, \boldsymbol{\psi}_h)\big\rangle = -\big(p_\mathrm{D}(t), \mathbf{n}\cdot\boldsymbol{\psi}_h\big)_{\Gamma_\mathrm{D}} - \big(g_\mathrm{N}(t), \varphi_h\big)_{\Gamma_\mathrm{N}} + \big(p_\mathrm{D}(t), Z_h^{-1}\varphi_h\big)_{\Gamma_\mathrm{D}} + \big(g_\mathrm{N}(t), Z_h\mathbf{n}\cdot\boldsymbol{\psi}_h\big)_{\Gamma_\mathrm{N}} . \tag{24}$$

Integration by parts gives

$$a_h\big((p_h, \mathbf{q}_h), (p_h, \mathbf{q}_h)\big) = \frac{1}{2}\sum_{K\in\mathcal{K}_h}\sum_{F\in\mathcal{F}_K} \frac{1}{Z_K + Z_{K_F}} \Bigg(\big\|[p_h]_{K,F}\big\|^2_F + Z_K Z_{K_F}\big\|\mathbf{n}_K\cdot[\mathbf{q}_h]_{K,F}\big\|^2_F \Bigg) .$$

Elastic waves. The full upwind DG approximation for the elastic wave equation (6) is given by

$$a_h\big((\mathbf{v}_h,\boldsymbol{\sigma}_h),(\mathbf{w}_h,\boldsymbol{\eta}_h)\big) = \sum_{K\in\mathcal{K}_h}\Bigg(\big(\boldsymbol{\sigma}_{h.K},\boldsymbol{\varepsilon}(\mathbf{w}_{h,K})\big)_K + \big(\mathbf{v}_{h,K},\nabla\cdot\boldsymbol{\eta}_{h,K}\big)_K \\ - \sum_{F\in\mathcal{F}_K}\frac{1}{Z^{\mathrm{p}}_K+Z^{\mathrm{p}}_{K_F}}\Big(\mathbf{n}_K\cdot\big(\boldsymbol{\sigma}_{h,K}\mathbf{n}_K - Z^{\mathrm{p}}_{K_F}\mathbf{v}_{h,K}\big),\mathbf{n}_K\cdot\big([\boldsymbol{\eta}_h]_{K,F}\mathbf{n}_K - Z^{\mathrm{p}}_K[\mathbf{w}_h]_{K,F}\big)\Big)_F \\ - \sum_{F\in\mathcal{F}_K}\frac{1}{Z^{\mathrm{s}}_K+Z^{\mathrm{s}}_{K_F}}\Big(\mathbf{n}_K\times\big(\boldsymbol{\sigma}_{h,K}\mathbf{n}_K - Z^{\mathrm{s}}_{K_F}\mathbf{v}_{h,K}\big),\mathbf{n}_K\times\big([\boldsymbol{\eta}_h]_{K,F}\mathbf{n}_K - Z^{\mathrm{s}}_K[\mathbf{w}_h]_{K,F}\big)\Big)_F\Bigg) \tag{25}$$

for $(\mathbf{v}_h,\boldsymbol{\sigma}_h),(\mathbf{w}_h,\boldsymbol{\eta}_h)\in\mathcal{S}_h$. The coefficients $Z^{\mathrm{p}}_K=\sqrt{(2\mu_{h,K}+\lambda_{h,K})\varrho_{h,K}}$ and $Z^{\mathrm{s}}_K=\sqrt{\mu_{h,K}\varrho_{h,K}}$ are the impedance of compressional waves and shear waves, respectively. On Dirichlet boundary faces $F\in\mathcal{F}_h\cap\Gamma_{\mathrm{D}}$, we set $[\mathbf{v}_h]_{K,F}=-2\mathbf{v}_h$ and $[\boldsymbol{\sigma}_h]_{K,F}\mathbf{n}_K=\mathbf{0}$, and on Neumann faces $F\in\mathcal{F}_h\cap\Gamma_{\mathrm{N}}$ we set $[\mathbf{v}_h]_{K,F}=\mathbf{0}$ and $[\boldsymbol{\sigma}_h]_{K,F}\mathbf{n}_K=-2\,\boldsymbol{\sigma}_h\mathbf{n}_K$. The right-hand side is given by

$$\big\langle\ell_{\partial\Omega,h}(t),(\mathbf{w}_h,\boldsymbol{\eta}_h)\big\rangle = \big(\mathbf{v}_{\mathrm{D}}(t),\boldsymbol{\eta}_h\mathbf{n}\big)_{\Gamma_{\mathrm{D}}} + \big(\mathbf{g}_{\mathrm{N}}(t),\mathbf{w}_h\big)_{\Gamma_{\mathrm{N}}} \\ + \big(\mathbf{n}\cdot\mathbf{v}_{\mathrm{D}}(t),(Z^{\mathrm{p}}_h)^{-1}\mathbf{n}\cdot\mathbf{w}_h\big)_{\Gamma_{\mathrm{D}}} + \big(\mathbf{n}\cdot\mathbf{g}_{\mathrm{N}}(t),Z^{\mathrm{p}}_h\mathbf{n}\cdot\boldsymbol{\eta}_h\mathbf{n}\big)_{\Gamma_{\mathrm{N}}} \\ + \big(\mathbf{n}\times\mathbf{v}_{\mathrm{D}}(t),(Z^{\mathrm{s}}_h)^{-1}\mathbf{n}\times\mathbf{w}_h\big)_{\Gamma_{\mathrm{D}}} + \big(\mathbf{n}\times\mathbf{g}_{\mathrm{N}}(t),Z^{\mathrm{s}}_h\mathbf{n}\times\boldsymbol{\eta}_h\mathbf{n}\big)_{\Gamma_{\mathrm{N}}}$$

with $Z^{\mathrm{p}}_h=Z^{\mathrm{p}}_K$ and $Z^{\mathrm{s}}_h=Z^{\mathrm{s}}_K$ on $\partial K\cap\partial\Omega$. Integrating by parts yields

$$a_h\big((\mathbf{v}_h,\boldsymbol{\sigma}_h),(\mathbf{v}_h,\boldsymbol{\sigma}_h)\big) = \frac{1}{2}\sum_{K\in\mathcal{K}_h}\sum_{F\in\mathcal{F}_K}\Bigg(\frac{\big\|\mathbf{n}_K\cdot\big([\boldsymbol{\sigma}_h]_{K,F}\mathbf{n}_K\big)\big\|_F^2 + Z^{\mathrm{p}}_K Z^{\mathrm{p}}_{K_F}\big\|\mathbf{n}_K\cdot[\mathbf{v}_h]_{K,F}\big\|_F^2}{Z^{\mathrm{p}}_K+Z^{\mathrm{p}}_{K_F}} \\ + \frac{\big\|\mathbf{n}_K\times\big([\boldsymbol{\sigma}_h]_{K,F}\mathbf{n}_K\big)\big\|_F^2 + Z^{\mathrm{s}}_K Z^{\mathrm{s}}_{K_F}\big\|\mathbf{n}_K\times[\mathbf{v}_h]_{K,F}\big\|_F^2}{Z^{\mathrm{s}}_K+Z^{\mathrm{s}}_{K_F}}\Bigg). \tag{26}$$

Electro-magnetic waves. The full upwind DG approximation for the electro-magnetic wave equation (7) is given by

$$a_h\big((\mathbf{E}_h,\mathbf{H}_h),(\boldsymbol{\varphi}_h,\boldsymbol{\psi}_h)\big) = \sum_{K\in\mathcal{K}_h}\Bigg(\big(\mathbf{E}_{h,K},\nabla\times\boldsymbol{\psi}_{h,K}\big)_K - \big(\mathbf{H}_{h,K},\nabla\times\boldsymbol{\varphi}_{h,K}\big)_K \\ + \sum_{F\in\mathcal{F}_K}\frac{1}{Z_K+Z_{K_F}}\Big(\big(Z_K\mathbf{E}_{h,K} - \mathbf{n}_K\times\mathbf{H}_{h,K},\mathbf{n}_K\times[\boldsymbol{\psi}_h]_{K,F}\big)_F \\ - \big(Z_K\mathbf{n}_K\times\mathbf{E}_{h,K} + \mathbf{H}_{h,K},Z_{K_F}\mathbf{n}_K\times[\boldsymbol{\varphi}_h]_{K,F}\big)_F\Big)\Bigg) \tag{27}$$

for $(\mathbf{E}_h,\mathbf{H}_h),(\boldsymbol{\varphi}_h,\boldsymbol{\psi}_h)\in\mathcal{S}_h$ with coefficient $Z_K=\sqrt{\varepsilon_K/\mu_K}$. On the boundary faces, we set $\mathbf{n}_K\times[\mathbf{E}]_{K,F}=-2\mathbf{n}_K\times\mathbf{E}_{h,K}$ and $\mathbf{n}_K\times[\mathbf{H}_h]_{K,F}=\mathbf{0}$ on $F\in\mathcal{F}_h\cap\Gamma_{\mathrm{E}}$, and on impedance boundary faces $F\in\mathcal{F}_h\cap\Gamma_{\mathrm{M}}$, we set $\mathbf{n}_K\times[\mathbf{E}]_{K,F}=\mathbf{0}$ and $\mathbf{n}_K\times[\mathbf{H}]_{K,F}=-2\mathbf{n}_K\times\mathbf{H}_{h,K}$. The right-hand side is given by

$$\big\langle\ell_{\partial\Omega,h}(t),(\boldsymbol{\varphi}_h,\boldsymbol{\psi}_h)\big\rangle = \big(\mathbf{g}_{\mathrm{M}}(t),\boldsymbol{\varphi}_h - Z_h^{-1}\mathbf{n}\times\boldsymbol{\psi}_h\big)_{\Gamma_{\mathrm{M}}}$$

with $Z_h=Z_K$ on $\partial K\cap\Gamma_{\mathrm{M}}$. Again, integration by parts yields

$$a_h\big((\mathbf{E}_h,\mathbf{H}_h),(\mathbf{E}_h,\mathbf{H}_h)\big) = \frac{1}{2}\sum_{K\in\mathcal{K}_h}\sum_{F\in\mathcal{F}_K}\frac{1}{Z_K+Z_{K_F}}\Big(Z_K Z_{K_F}\big\|\mathbf{n}_K\times[\mathbf{E}_h]_{K,F}\big\|_F^2 + \big\|\mathbf{n}_K\times[\mathbf{H}_h]_{K,F}\big\|_F^2\Big)\,.$$

3.4. A discontinuous Galerkin method in time and space

Combining the two semi-discretizations, we obtain the full DG discretization

$$b_h(\mathbf{v}_h, \mathbf{w}_h) = m_h(\mathbf{v}_h, \mathbf{w}_h) + \int_0^T a_h\big(\mathbf{v}_h(t), \mathbf{w}_h(t)\big)\,\mathrm{d}t\,, \qquad \mathbf{v}_h\,,\ \mathbf{w}_h \in \mathcal{V}_h \tag{28}$$

with right-hand side in the space-time cylinder

$$\big\langle \ell_h, \mathbf{w}_h\big\rangle = \big(\mathbf{f}, \mathbf{w}_h\big)_Q + \big(M_h\mathbf{u}_0, \mathbf{w}_h(0)\big)_\Omega + \int_0^T \big\langle \ell_{\partial\Omega,h}(t), \mathbf{w}_h(t)\big\rangle\,\mathrm{d}t\,, \qquad \mathbf{v}_h \in \mathcal{V}_h\,. \tag{29}$$

For the space-time DG method we have by construction dual consistency for the bilinear form and the right hand side

$$b_h\big(\mathbf{v}_h, \mathbf{w}\big) = \big(\mathbf{v}_h, L_h^*\mathbf{w}\big)_{Q_h} \quad\text{and}\quad \big\langle \ell_h, \mathbf{w}\big\rangle = \big(\mathbf{f}, \mathbf{w}\big)_Q + \big(M_h\mathbf{u}_0, \mathbf{w}(0)\big)_{\partial\Omega} + \big(\mathbf{g}, \mathbf{w}\big)_{(0,T)\times\partial\Omega}\,, \quad \mathbf{v}_h \in \mathcal{V}_h\,,\ \mathbf{w} \in \mathcal{V}^*\,, \tag{30}$$

and positivity for the inconsistency complement

$$b_h\big(\mathbf{v}_h, \mathbf{v}_h\big) \ge \frac{1}{2}\sum_{n=0}^{N} \big\|M_h^{1/2}[\mathbf{v}_h]_n\big\|_\Omega^2 + c_1\big\|\underline{A}_{\mathbf{n}}[\mathbf{v}_h]\big\|_{\partial\Omega_h}^2\,, \qquad \mathbf{v}_h \in \mathcal{V}_h \tag{31}$$

by combining (17) and (20). Together with (18) and (21) we obtain

$$b_h\big(\mathbf{v}_h, \mathbf{v}_h\big) = 0 \quad\Longrightarrow\quad b_h\big(\mathbf{v}_h, \mathbf{w}_h\big) = \big(\mathbf{v}_h, L_h^*\mathbf{w}_h\big)_{Q_h} = \big(L_h\mathbf{v}_h, \mathbf{w}_h\big)_{Q_h}\,, \qquad \mathbf{v}_h, \mathbf{w}_h \in \mathcal{V}_h\,, \tag{32}$$

and (22) yields with $C_1 > 0$

$$\big|b_h\big(\mathbf{v}_h, \mathbf{w}_h\big) - \big(\mathbf{v}_h, L_h^*\mathbf{w}_h\big)_{\Omega_h}\big| \le \big\|M_h^{1/2}\mathbf{v}_h\big\|_{\partial Q_h}\sqrt{\big\|M_h^{1/2}[\mathbf{w}_h]\big\|_{\partial I_h\times\Omega}^2 + C_1\big\|\underline{A}_{\mathbf{n}}[\mathbf{w}_h]\big\|_{\partial\Omega_h}^2}\,, \qquad \mathbf{v}_h, \mathbf{w}_h \in \mathcal{V}_h\,, \tag{33a}$$

$$\big|b_h\big(\mathbf{v}_h, \mathbf{w}_h\big) - \big(L_h\mathbf{v}_h, \mathbf{w}_h\big)_{\Omega_h}\big| \le \sqrt{\big\|M_h^{1/2}[\mathbf{v}_h]\big\|_{\partial I_h\times\Omega}^2 + C_1\big\|\underline{A}_{\mathbf{n}}[\mathbf{v}_h]\big\|_{\partial\Omega_h}^2}\;\big\|M_h^{1/2}\mathbf{w}_h\big\|_{\partial Q_h}\,, \qquad \mathbf{v}_h, \mathbf{w}_h \in \mathcal{V}_h\,, \tag{33b}$$

$$\big|\big\langle \ell, \mathbf{w}_h\big\rangle - \big\langle \ell_h, \mathbf{w}_h\big\rangle\big| \le \big\|M_h^{1/2}\mathbf{u}_0\big\|_\Omega\big\|M_h^{1/2}\mathbf{w}_h\big\|_\Omega + C_1\big\|\mathbf{g}\big\|_{I_h\times\partial\Omega}\big\|M_h^{1/2}\mathbf{w}_h\big\|_{I_h\times\partial\Omega}\,, \qquad \mathbf{w}_h \in \mathcal{V}_h\,. \tag{33c}$$

For sufficiently smooth functions $\mathbf{v} \in \mathrm{L}_2(Q;\mathbb{R}^m)$ with $L_h\mathbf{v} \in \mathrm{L}_2(Q;\mathbb{R}^m)$, $\mathbf{v}(0) \in \mathrm{L}_2(\Omega;\mathbb{R}^m)$, $[\mathbf{v}]_n = \mathbf{0}$ for $n = 1, \dots, N-1$, $A_{\mathbf{n}}[\mathbf{v}] = \mathbf{0}$ on $I_h \times F$ for inner faces $F \in \mathcal{F}_h \setminus \partial\Omega$, and $A_{\mathbf{n}}[\mathbf{v}] \in \mathrm{L}_2(I \times \partial\Omega;\mathbb{R}^m)$, we obtain consistency of the form

$$\big|b_h\big(\mathbf{v}, \mathbf{w}_h\big) - \big\langle \ell_h, \mathbf{w}_h\big\rangle - \big(L_h\mathbf{v} - \mathbf{f}, \mathbf{w}_h\big)_Q - \big(M_h(\mathbf{v}(0) - \mathbf{u}_0), \mathbf{w}_h\big)_\Omega\big| \le C_1\sum_{k=1}^{m}\big\|(A_{\mathbf{n}}\mathbf{v})_k - g_k\big\|_{I_h\times\Gamma_k}\big\|w_{h,k}\big\|_{I_h\times\Gamma_k}\,. \tag{34}$$

Lemma 7. *We have, depending on $c_1 > 0$ in (20),*

$$\big\|M_h^{1/2}\mathbf{v}_h\big\|_Q^2 + \sum_{n=0}^{N-1} d_T(t_n)\big\|M_h^{1/2}[\mathbf{v}_h]_n\big\|_\Omega^2 + 2c_1\int_0^T d_T(t)\big\|\underline{A}_{\mathbf{n}}[\mathbf{v}_h(t)]\big\|_{\partial\Omega_h}^2\,\mathrm{d}t \le 2\,b_h(\mathbf{v}_h, d_T\mathbf{v}_h)\,, \qquad \mathbf{v}_h \in \mathcal{V}_h\,.$$

Proof. By inserting $\mathbf{v}_h(t)$ into (20) and integrating over time we find

$$c_1\int_0^T d_T(t)\big\|\underline{A}_{\mathbf{n}}[\mathbf{v}_h(t)]\big\|_{\partial\Omega_h}^2\,\mathrm{d}t \le \int_0^T d_T(t)a_h(\mathbf{v}_h(t), \mathbf{v}_h(t))\,\mathrm{d}t\,,$$

and thus with Lem. 6 we get for all $\mathbf{v}_h \in \mathcal{V}_h$

$$\begin{aligned}
\big(M_h\mathbf{v}_h, \mathbf{v}_h\big)_Q + \sum_{n=0}^{N-1} d_T(t_n)\big(M_h[\mathbf{v}_h]_n, [\mathbf{v}_h]_n\big)_\Omega &\le 2\,m_h(\mathbf{v}_h, d_T\mathbf{v}_h)\\
&\le 2\,m_h(\mathbf{v}_h, d_T\mathbf{v}_h) + 2\int_0^T d_T(t)\,a_h\big(\mathbf{v}_h(t), \mathbf{v}_h(t)\big)\,\mathrm{d}t - 2c_1\int_0^T d_T(t)\big\|\underline{A}_{\mathbf{n}}[\mathbf{v}_h(t)]\big\|_{\partial\Omega_h}^2\,\mathrm{d}t\\
&= 2\,b_h(\mathbf{v}_h, d_T\mathbf{v}_h) - 2c_1\int_0^T d_T(t)\big\|\underline{A}_{\mathbf{n}}[\mathbf{v}_h(t)]\big\|_{\partial\Omega_h}^2\,\mathrm{d}t\,.
\end{aligned}$$

□

4. Well-posedness and stability

We show that the discrete problem has a unique solution and is stable with respect to different norms.

4.1. Well-posedness of the space-time DG discretization

The well-posedness of the discrete equation is now established as in [Bansal et al., 2021, Prop. 5.1].

Lemma 8. *A unique discrete approximation* $\mathbf{u}_h \in V_h$ *exists solving*

$$b_h(\mathbf{u}_h, \mathbf{v}_h) = \langle \ell_h, \mathbf{v}_h \rangle\,, \qquad \mathbf{v}_h \in V_h\,. \tag{35}$$

Proof. Since $\dim V_h < \infty$, it is sufficient to show that $\mathbf{u}_h = \mathbf{0}$ is the unique solution of the homogeneous problem

$$b_h(\mathbf{u}_h, \mathbf{v}_h) = 0\,, \qquad \mathbf{v}_h \in V_h\,. \tag{36}$$

Since (36) implies $b_h(\mathbf{u}_h, \mathbf{u}_h) = 0$, we obtain by (31) for the jump terms $\big\|M_h^{1/2}[\mathbf{u}_h]\big\|_{\partial I_h \times \Omega_h} = \big\|\underline{A}_{\mathbf{n}}[\mathbf{u}_h]\big\|_{I_h \times \partial\Omega_h} = 0$, so that $b_h(\mathbf{u}_h, \mathbf{v}_h) = \big(L_h\mathbf{u}_h, \mathbf{v}_h\big)_{Q_h} = 0$. Since M_h is piecewise constant in $K \in \mathcal{K}_h$, we observe $L_h\mathbf{u}_h \in V_h$, so that we can test with $\mathbf{v}_h = L_h\mathbf{u}_h$; thus, also $\big(L_h\mathbf{u}_h, L_h\mathbf{u}_h\big)_{Q_h} = 0$, i.e., $L_h\mathbf{u}_h = \mathbf{0}$. Now the assertion follows from Lem. 7 and (32) by

$$\big\|M_h^{1/2}\mathbf{u}_h\big\|_Q^2 = \big(M_h\mathbf{u}_h, \mathbf{u}_h\big)_Q \le 2\, b_h(\mathbf{u}_h, d_T\mathbf{u}_h) = 2\,\big(L_h\mathbf{u}_h, d_T\mathbf{u}_h\big)_Q = 0\,.$$

□

Remark 9. *The previous lemma shows that the discrete graph norm defined by*

$$\|\mathbf{v}_h\|_{V_h} = \sup_{\mathbf{w}_h \in V_h \setminus \{\mathbf{0}\}} \frac{b_h(\mathbf{v}_h, \mathbf{w}_h)}{\big\|M_h^{1/2}\mathbf{w}_h\big\|_Q}\,, \qquad \mathbf{v}_h \in V_h\,, \tag{37}$$

is well defined and a norm in V_h.
Since the discrete graph norm is only a semi-norm in $\mathcal{V}_h$, *we have to use stronger norms for the convergence analysis.*

4.2. Stability in space and time

Let $0 = c_{p,0} < c_{p,1} < \cdots < c_{p,p} < 1$ be the Radau Ia collocation points, so that

$$\int_0^1 \phi(s)\,\mathrm{d}s = \sum_{k=0}^{p} \omega_{p,k}\phi(c_{p,k})\,, \qquad \phi \in \mathbb{P}_{2p}$$

(with quadrature weights $\omega_{p,k} > 0$ for $k = 0, \dots, p$), and let $\lambda_{p,k} \in \mathbb{P}_p$ be the corresponding Lagrange polynomials

$$\lambda_{p,k}(s) = \prod_{j=0,\ j\neq k}^{p} \frac{s - c_{p,j}}{c_{p,k} - c_{p,j}}\,, \qquad s \in [0,1]\,.$$

This defines $\lambda_{n,h,k} \in \mathbb{P}_{p_n}(I_{n,h})$ by $\lambda_{n,h,k}(t_{n-1} + s\triangle t_n) = \lambda_{p_n,k}(s)$ for $s \in [0,1]$ and $t_{n,k} = t_{n-1} + c_{p_n,k}\triangle t_n$.
Together this is combined to the corresponding interpolation $\mathcal{I}_h \colon \mathcal{V}_h \longrightarrow V_h$ by

$$(\mathcal{I}_{n,h}\mathbf{v}_{n,h})(t,\mathbf{x}) = \sum_{k=0}^{p_n} \lambda_{n,h,k}(t)\mathbf{v}_{n,h}(t_{n,k},\mathbf{x})\,, \quad (t,\mathbf{x}) \in I_{n,h} \times \Omega_h,\ \ \mathbf{v}_{n,h} \in \mathrm{C}^0([t_{n-1},t_n];\mathcal{S}_h),\ \ n = 1,\dots,N\,.$$

For the interpolation we will use in the following the estimate

$$\begin{aligned}
\big\|M_h^{1/2}\mathcal{I}_h(d_T\mathbf{v}_h)\big\|_Q^2 &= \sum_{n=1}^{N}\sum_{k=0}^{p_n} \omega_{p_n,k}\big\|M_h^{1/2}\mathcal{I}_h(d_T\mathbf{v}_h)(t_{n,k})\big\|_\Omega^2 = \sum_{n=1}^{N}\sum_{k=0}^{p_n} d_T(t_{n,k})^2\omega_{p_n,k}\big\|M_h^{1/2}\mathbf{v}_h(t_{n,k})\big\|_\Omega^2 \\
&\le T^2 \sum_{n=1}^{N}\sum_{k=0}^{p_n} \omega_{p_n,k}\big\|M_h^{1/2}\mathbf{v}_h(t_{n,k})\big\|_\Omega^2 = T^2\,\big\|M_h^{1/2}\mathbf{v}_h\big\|_Q^2\,.
\end{aligned} \tag{38}$$

Lemma 10. *If $p_{n,K} = p_n$ for all $K \in \mathcal{K}_h$ and $n = 1, \dots, N$, we have for $\mathbf{v}_h \in V_h$*

$$\big\|M_h^{1/2}\mathbf{v}_h\big\|_Q^2 + \sum_{n=1}^N \left(d_T(t_{n-1})\big\|M_h^{1/2}[\mathbf{v}_h]_{n-1}\big\|_\Omega^2 + 2c_1 \sum_{k=0}^{p_n} d_T(t_{n,k})\omega_{p_n,k}\big\|\underline{A}_{\mathbf{n}}[\mathbf{v}_h(t_{n,k})]\big\|_{\partial\Omega_h}^2 \right) \le 2\, b_h\big(\mathbf{v}_h, \mathcal{I}_h(d_T\mathbf{v}_h)\big)\,.$$

Proof. We observe

$$\begin{aligned}\big(M_h\partial_t\mathbf{v}_h, d_T\mathbf{v}_h\big)_{Q_h} &= \sum_{n=1}^N \big(M_h\partial_t\mathbf{v}_{n,h}, d_T\mathbf{v}_{n,h}\big)_{I_{n,h}\times\Omega} = \sum_{n=1}^N\sum_{k=0}^{p_n} \omega_{p_n,k}\big(M_h(\partial_t\mathbf{v}_{n,h})(t_{n,k}), d_T(t_{n,k})\mathbf{v}_{n,h}(t_{n,k})\big)_\Omega \\ &= \sum_{n=1}^N\sum_{k=0}^{p_n} \omega_{p_n,k}\big(M_h(\partial_t\mathbf{v}_{n,h})(t_{n,k}), \mathcal{I}_{n,h}(d_T\mathbf{v}_{n,h})(t_{n,k})\big)_\Omega = \big(M_h\partial_t\mathbf{v}_h, \mathcal{I}_h(d_T\mathbf{v}_h)\big)_{Q_h}\,.\end{aligned}$$

Using $\mathcal{I}_h(d_T\mathbf{v}_h)(t_{n-1}) = d_T(t_{n-1})\mathbf{v}_{n,h}(t_{n-1})$ for $n = 1, \dots, N$, we have

$$\begin{aligned}m_h(\mathbf{v}_h, d_T\mathbf{v}_h) &= \big(M_h\partial_t\mathbf{v}_h, d_T\mathbf{v}_h\big)_{Q_h} + \sum_{n=1}^N \big(M_h[\mathbf{v}_h]_n, d_T(t_{n-1})\mathbf{v}_{n,h}(t_{n-1})\big)_\Omega \\ &= \big(M_h\partial_t\mathbf{v}_h, \mathcal{I}_h(d_T\mathbf{v}_h)\big)_{Q_h} + \sum_{n=1}^N \big(M_h[\mathbf{v}_h]_n, \mathcal{I}_h(d_T\mathbf{v}_h)(t_{n-1})\big)_\Omega = m_h\big(\mathbf{v}_h, \mathcal{I}_h(d_T\mathbf{v}_h)\big)\,,\end{aligned}$$

and together with Lem. 6 we obtain

$$\big\|M_h^{1/2}\mathbf{v}_h\big\|_Q^2 + \sum_{n=1}^N d_T(t_{n-1})\big\|M_h^{1/2}[\mathbf{v}_h]_{n-1}\big\|_\Omega^2 \le 2\, m_h(\mathbf{v}_h, d_T\mathbf{v}_h) = 2\, m_h\big(\mathbf{v}_h, \mathcal{I}_h(d_T\mathbf{v}_h)\big)\,.$$

For the upwind DG discretization in space we obtain by (20)

$$\begin{aligned}0 &\le c_1 \sum_{n=1}^N\sum_{k=0}^{p_n} d_T(t_{n,k})\omega_{p_n,k}\big\|\underline{A}_{\mathbf{n}}[\mathbf{v}_h(t_{n,k})]\big\|_{\partial\Omega_h}^2 \\ &\le \sum_{n=1}^N\sum_{k=0}^{p_n} d_T(t_{n,k})\omega_{p_n,k}\, a_h\big(\mathbf{v}_h(t_{n,k}), \mathbf{v}_h(t_{n,k})\big) = \sum_{n=1}^N\sum_{k=0}^{p_n} \omega_{p_n,k}\, a_h\big(\mathbf{v}_h(t_{n,k}), \mathcal{I}_{n,h}(d_T\mathbf{v}_{n,h})(t_{n,k})\big) \\ &= \sum_{n=1}^N \int_{t_{n-1}}^{t_n} a_h\big(\mathbf{v}_h(t), \mathcal{I}_{n,h}(d_T\mathbf{v}_{n,h})(t)\big)\,\mathrm{d}t = \int_0^T a_h\big(\mathbf{v}_h(t), \mathcal{I}_h(d_T\mathbf{v}_h)(t)\big)\,\mathrm{d}t\,,\end{aligned}$$

so that together we obtain the assertion by

$$\begin{aligned}&\big\|M_h^{1/2}\mathbf{v}_h\big\|_Q^2 + \sum_{n=1}^N \left(d_T(t_{n-1})\big\|M_h^{1/2}[\mathbf{v}_h]_{n-1}\big\|_\Omega^2 + 2c_1 \sum_{k=0}^{p_n} d_T(t_{n,k})\omega_{p_n,k}\big\|\underline{A}_{\mathbf{n}}[\mathbf{v}_h(t_{n,k})]\big\|_{\partial\Omega_h}^2 \right) \\ &\qquad\le 2\, m_h\big(\mathbf{v}_h, \mathcal{I}_h(d_T\mathbf{v}_h)\big) + \int_0^T a_h\big(\mathbf{v}_h(t), \mathcal{I}_h(d_T\mathbf{v}_h)(t)\big)\,\mathrm{d}t = 2\, b_h\big(\mathbf{v}_h, \mathcal{I}_h(d_T\mathbf{v}_h)\big)\,.\end{aligned}$$

□

Remark 11. *Together with (38) we obtain L_2 stability with respect to the discrete graph norm by*

$$\big\|M_h^{1/2}\mathbf{v}_h\big\|_Q \le 2\frac{b_h\big(\mathbf{v}_h, \mathcal{I}_h(d_T\mathbf{v}_h)\big)}{\big\|\mathcal{I}_h(d_T\mathbf{v}_h)\big\|_Q}\,\frac{\big\|\mathcal{I}_h(d_T\mathbf{v}_h)\big\|_Q}{\big\|M_h^{1/2}\mathbf{v}_h\big\|_Q} \le 2T\,\|\mathbf{v}_h\|_{V_h}$$

for $\mathbf{v}_h \in V_h \setminus \{0\}$, i.e., $\|M_h^{1/2}\mathbf{v}_h\|_Q \le 2T\,\|\mathbf{v}_h\|_{V_h}$.

Corollary 12. *Let $\mathbf{u}_h \in V_h$ be the discrete solution (35), and assume homogeneous boundary data $\mathbf{g} = \mathbf{0}$. If $p_{n,K} = p_n$ for all $K \in \mathcal{K}_h$ and $n = 1, \dots, N$, the solution is bounded by*

$$\big\|M_h^{1/2}\mathbf{u}_h\big\|_Q^2 + \sum_{n=1}^N d_T(t_{n-1}) \Big(\big\|M_h^{1/2}[\mathbf{u}_h]_{n-1}\big\|_\Omega^2 + 2c_1\big\|\underline{A}_\mathbf{n}[\mathbf{u}_h]\big\|_{I_{n,h}\times\partial\Omega_h}^2\Big) \le 4\,\big\|d_T M_h^{-1/2}\mathbf{f}\big\|_Q^2 + 4T\,\big\|M_h^{1/2}\mathbf{u}_0\big\|_\Omega^2 .$$

Proof. We have $\mathcal{I}_h(d_T\mathbf{u}_h)(0) = \mathbf{u}_h(0)$ and for $n = 1, \dots, N$

$$d_T(t_{n-1})\big\|\underline{A}_\mathbf{n}[\mathbf{u}_h]\big\|_{I_{n,h}\times\partial\Omega_h}^2 = d_T(t_{n-1}) \sum_{k=0}^{p_n} \omega_{p_n,k}\big\|\underline{A}_\mathbf{n}[\mathbf{u}_h(t_{n,k})]\big\|_{\partial\Omega_h}^2 \le \sum_{k=0}^{p_n} d_T(t_{n,k})\omega_{p_n,k}\big\|\underline{A}_\mathbf{n}[\mathbf{u}_h(t_{n,k})]\big\|_{\partial\Omega_h}^2 ,$$

so that together with Lem. 10 we get the assertion by

$$\begin{aligned} \frac{1}{2}\big\|M_h^{1/2}\mathbf{u}_h\big\|_Q^2 &+ \frac{1}{2}\sum_{n=1}^N d_T(t_{n-1}) \Big(\big\|M_h^{1/2}[\mathbf{u}_h]_{n-1}\big\|_\Omega^2 + 2c_1\big\|\underline{A}_\mathbf{n}[\mathbf{u}_h]\big\|_{I_{n,h}\times\partial\Omega_h}^2\Big) \\ &\le b_h\big(\mathbf{u}_h, \mathcal{I}_h(d_T\mathbf{u}_h)\big) = \langle \ell_h, \mathcal{I}_h(d_T\mathbf{u}_h)\rangle = \big(\mathbf{f}, d_T\mathbf{u}_h\big)_Q + \big(M_h\mathbf{u}_0, T\mathbf{u}_h(0)\big)_\Omega \\ &\le \big\|d_T M_h^{-1/2}\mathbf{f}\big\|_Q^2 + \frac{1}{4}\big\|M_h^{1/2}\mathbf{u}_h\big\|_Q^2 + T\big\|M_h^{1/2}\mathbf{u}_0\big\|_\Omega^2 + \frac{T}{4}\big\|M_h^{1/2}\mathbf{u}_h(0)\big\|_\Omega^2 . \end{aligned}$$

□

Remark 13. *The estimate in Lem. 7 directly implies that the Petrov–Galerkin method with test space $V_h^* = d_T V_h$ is well-defined and L_2 stable: the Petrov–Galerkin solution $\mathbf{u}_h^{\mathrm{PG}} \in V_h$ given by*

$$b_h(\mathbf{u}_h^{\mathrm{PG}}, d_T\mathbf{v}_h) = \langle \ell_h, d_T\mathbf{v}_h\rangle , \qquad \mathbf{v}_h \in V_h \tag{39}$$

is bounded by

$$\frac{1}{2}\big\|M_h^{1/2}\mathbf{u}_h^{\mathrm{PG}}\big\|_Q^2 + \frac{T}{2}\big\|M_h^{1/2}\mathbf{u}_h^{\mathrm{PG}}(0)\big\|_\Omega^2 \le b_h(\mathbf{u}_h^{\mathrm{PG}}, d_T\mathbf{u}_h^{\mathrm{PG}}) = \langle \ell_h, d_T\mathbf{u}_h^{\mathrm{PG}}\rangle ,$$

and thus, in case of homogeneous boundary data $\mathbf{g} = \mathbf{0}$ we obtain

$$\big\|M_h^{1/2}\mathbf{u}_h^{\mathrm{PG}}\big\|_Q^2 + T\,\big\|M_h^{1/2}\mathbf{u}_h^{\mathrm{PG}}(0)\big\|_\Omega^2 \le 4\,\big\|d_T M_h^{-1/2}\mathbf{f}\big\|_Q^2 + 4T\,\big\|M_h^{1/2}\mathbf{u}_0\big\|_\Omega^2 .$$

This is proposed and analyzed in [Babuška et al., 2001] for the semi-discrete case. Our numerical tests indicate, that the Petrov–Galerkin modification does not improve the approximation quality, and in the next section we show, that stability and convergence in the DG norm can be established also for the Galerkin method with ansatz and test space V_h and with adaptively chosen $p_{n,K}$.

4.3. Inf-sup stability in the DG norm

Suitable mesh-dependent DG semi-norms and norms can be defined for $\mathbf{v}_h \in \mathcal{V}_h$ by

$$\big|\mathbf{v}_h\big|_{h,\mathrm{DG}} = \sqrt{b_h(\mathbf{v}_h,\mathbf{v}_h)}\,, \qquad \big|\mathbf{v}_h\big|_{h,\mathrm{DG}^+} = \sqrt{\sum_{n=1}^N \Big(\big\|M_h^{1/2}\mathbf{v}_{n,h}(t_{n-1})\big\|_\Omega^2 + \big\|M_h^{1/2}\mathbf{v}_{n,h}(t_n)\big\|_\Omega^2\Big) + C_1\big\|M_h^{1/2}\mathbf{v}_h\big\|_{I_h\times\partial\Omega_h}^2}\,,$$
$$\big\|\mathbf{v}_h\big\|_{h,\mathrm{DG}} = \sqrt{\big|\mathbf{v}_h\big|_{h,\mathrm{DG}}^2 + \big\|h^{1/2}\,M_h^{-1/2}L_h\mathbf{v}_h\big\|_{Q_h}^2}\,, \quad \big\|\mathbf{v}_h\big\|_{h,\mathrm{DG}^+} = \sqrt{\big|\mathbf{v}_h\big|_{h,\mathrm{DG}^+}^2 + \big\|h^{-1/2}M_h^{1/2}\mathbf{v}_h\big\|_Q^2}\,, \tag{40}$$

see [Ern and Guermond, 2021, Chap. 57], [Di Pietro and Ern, 2011, Chap. 2 and 7]. Analogously to the proof of Lem. 8 we observe that $\|\mathbf{v}_h\|_{h,\mathrm{DG}} = 0$ implies $\mathbf{v}_h = \mathbf{0}$, so that $\|\cdot\|_{h,\mathrm{DG}}$ indeed is a norm. Using (33), we obtain for $\mathbf{v}_h,\mathbf{w}_h \in \mathcal{V}_h$

$$\big|b_h(\mathbf{v}_h,\mathbf{w}_h) + \big(\mathbf{v}_h, L_h\mathbf{w}_h\big)_{Q_h}\big| \le \big|\mathbf{v}_h\big|_{h,\mathrm{DG}^+}\big|\mathbf{w}_h\big|_{h,\mathrm{DG}} \quad \text{and} \quad \big|b_h(\mathbf{v}_h,\mathbf{w}_h) - \big(L_h\mathbf{v}_h, \mathbf{w}_h\big)_{Q_h}\big| \le \big|\mathbf{v}_h\big|_{h,\mathrm{DG}}\big|\mathbf{w}_h\big|_{h,\mathrm{DG}^+}\,. \tag{41}$$

We have $2\,\big|\mathbf{v}_h\big|_{h,\mathrm{DG}}^2 = 2\,b_h(\mathbf{v}_h,\mathbf{v}_h) + \big(L_h\mathbf{v}_h,\mathbf{v}_h\big)_{Q_h} - \big(\mathbf{v}_h,L_h\mathbf{v}_h\big)_{Q_h} \le 2\,\big|\mathbf{v}_h\big|_{h,\mathrm{DG}^+}\big|\mathbf{v}_h\big|_{h,\mathrm{DG}}$, i.e., $\big|\mathbf{v}_h\big|_{h,\mathrm{DG}} \le \big|\mathbf{v}_h\big|_{h,\mathrm{DG}^+}$, and continuity of the bilinear form $b_h(\mathbf{v}_h,\mathbf{w}_h) \le \big\|\mathbf{v}_h\big\|_{h,\mathrm{DG}}\big\|\mathbf{w}_h\big\|_{h,\mathrm{DG}^+}$ and $b_h(\mathbf{v}_h,\mathbf{w}_h) \le \big\|\mathbf{v}_h\big\|_{h,\mathrm{DG}^+}\big\|\mathbf{w}_h\big\|_{h,\mathrm{DG}}$.

The inf-sup stability for the advection equation [Di Pietro and Ern, 2011, Lem. 2.35] can be transferred to our setting.

Theorem 14. *A constant $c_{\text{inf-sup}} > 0$ exists such that*

$$\sup_{\mathbf{w}_h\in V_h\setminus\{\mathbf{0}\}} \frac{b_h(\mathbf{v}_h,\mathbf{w}_h)}{\big\|\mathbf{w}_h\big\|_{h,\mathrm{DG}}} \ge c_{\text{inf-sup}}\big\|\mathbf{v}_h\big\|_{h,\mathrm{DG}}\,, \qquad \mathbf{v}_h \in V_h\,.$$

Proof. For given $\mathbf{v}_h \in V_h \setminus \{\mathbf{0}\}$ we define $\mathbf{z}_h = hM_h^{-1}L_h\mathbf{v}_h \in V_h$, and we obtain by the discrete trace inequality (12b)

$$\big|\mathbf{z}_h\big|_{h,\mathrm{DG}^+} \le C_{\mathrm{tr}}\big\|h^{-1/2}M_h^{1/2}\mathbf{z}_h\big\|_{Q_h} = C_{\mathrm{tr}}\big\|h^{1/2}M_h^{-1/2}L_h\mathbf{v}_h\big\|_{Q_h} \le C_{\mathrm{tr}}\big\|\mathbf{v}_h\big\|_{h,\mathrm{DG}}\,, \tag{42}$$

and together with the inverse inequality (12a) this yields

$$\big\|\mathbf{z}_h\big\|_{h,\mathrm{DG}}^2 = \big|\mathbf{z}_h\big|_{h,\mathrm{DG}}^2 + \big\|h^{1/2}M_h^{-1/2}L_h\mathbf{z}_h\big\|_{Q_h}^2 \le \big|\mathbf{z}_h\big|_{h,\mathrm{DG}^+}^2 + C_{\mathrm{inv}}^2\big\|h^{-1/2}M_h^{1/2}\mathbf{z}_h\big\|_{Q_h}^2 \le \big(C_{\mathrm{tr}}^2 + C_{\mathrm{inv}}^2\big)\big\|\mathbf{v}_h\big\|_{h,\mathrm{DG}}^2\,. \tag{43}$$

We observe, using (41),

$$\big(L_h\mathbf{v}_h,\mathbf{z}_h\big)_{Q_h} - b_h(\mathbf{v}_h,\mathbf{z}_h) \le \big|\mathbf{v}_h\big|_{h,\mathrm{DG}}\big|\mathbf{z}_h\big|_{h,\mathrm{DG}^+} \le \frac{C_{\mathrm{tr}}^2}{2}\big|\mathbf{v}_h\big|_{h,\mathrm{DG}}^2 + \frac{1}{2C_{\mathrm{tr}}^2}\big|\mathbf{z}_h\big|_{h,\mathrm{DG}^+}^2 \le \frac{C_{\mathrm{tr}}^2}{2}\big|\mathbf{v}_h\big|_{h,\mathrm{DG}}^2 + \frac{1}{2}\big\|\mathbf{v}_h\big\|_{h,\mathrm{DG}}^2\,.$$

This yields, inserting $\big\|h^{1/2}M_h^{-1/2}L_h\mathbf{v}_h\big\|_{Q_h}^2 = \big(L_h\mathbf{v}_h,\mathbf{z}_h\big)_{Q_h}$,

$$\big\|\mathbf{v}_h\big\|_{h,\mathrm{DG}}^2 = \big|\mathbf{v}_h\big|_{h,\mathrm{DG}}^2 + \big(L_h\mathbf{v}_h,\mathbf{z}_h\big)_{Q_h} \le \big|\mathbf{v}_h\big|_{h,\mathrm{DG}}^2 + \frac{C_{\mathrm{tr}}^2}{2}\big|\mathbf{v}_h\big|_{h,\mathrm{DG}}^2 + \frac{1}{2}\big\|\mathbf{v}_h\big\|_{h,\mathrm{DG}}^2 + b_h(\mathbf{v}_h,\mathbf{z}_h)\,,$$

so that with $C_2 = 2 + C_{\mathrm{tr}}^2$

$$\big\|\mathbf{v}_h\big\|_{h,\mathrm{DG}}^2 \le C_2\big|\mathbf{v}_h\big|_{h,\mathrm{DG}}^2 + 2\,b_h(\mathbf{v}_h,\mathbf{z}_h) = b_h(\mathbf{v}_h, C_2\mathbf{v}_h + 2\mathbf{z}_h)\,. \tag{44}$$

Using (43), we obtain the assertion with $c_{\text{inf-sup}} = \Big(C_2 + 2\sqrt{C_{\mathrm{tr}}^2 + C_{\mathrm{inv}}^2}\Big)^{-1}$ by

$$\big\|\mathbf{v}_h\big\|_{h,\mathrm{DG}}^2 \le \big\|C_2\mathbf{v}_h + 2\mathbf{z}_h\big\|_{h,\mathrm{DG}} \frac{b_h(\mathbf{v}_h, C_2\mathbf{v}_h + 2\mathbf{z}_h)}{\big\|C_2\mathbf{v}_h + 2\mathbf{z}_h\big\|_{h,\mathrm{DG}}} \le c_{\text{inf-sup}}^{-1}\big\|\mathbf{v}_h\big\|_{h,\mathrm{DG}} \sup_{\mathbf{w}_h\in V_h\setminus\{\mathbf{0}\}} \frac{b_h(\mathbf{v}_h,\mathbf{w}_h)}{\big\|\mathbf{w}_h\big\|_{h,\mathrm{DG}}}\,.$$

□

5. Convergence of the DG space-time approximation

In the first step, we show that stability in L_2 implies convergence in the limit of the DG approximation. Then, by assuming some regularity of the solution, qualitative convergence results are obtained in the DG norm.

5.1. Convergence in the limit

Let $\big(Q_h\big)_{h\in\mathcal{H}}$ be a shape-regular family of space-time meshes with $\mathcal{H}=\{h_0,h_1,h_2,\cdots\}\subset(0,\infty)$ and $0\in\overline{\mathcal{H}}$.

Let $\big(V_h\big)_{h\in\mathcal{H}}$ be corresponding DG finite element spaces, so that

$$\lim_{h\in\mathcal{H}}\inf_{\mathbf{v}_h\in V_h}\big\|\mathbf{v}-\mathbf{v}_h\big\|_Q=0\,,\qquad \mathbf{v}\in\mathcal{V}^*\,. \tag{45}$$

For $h\in\mathcal{H}$, let $\mathbf{u}_h\in V_h$ be the solution of the discrete problem (35).

The proof of existence of a unique discrete solution in Lem. 8 only relies on the properties (31) and (32) of the DG bilinear form and thus only implicitly on the boundary parts $\Gamma_k\subset\partial\Omega$. In order to obtain a unique weak solution of (2) in the limit, constraints for the selection of $\Gamma_k\subset\partial\Omega$, $k=1,\dots,m$, are necessary, cf. (8). This is used in the following.

Theorem 15. *Assume that $p_{n,K}=p_n\ge1$ and $q_{n,K}\ge1$. In case of homogeneous boundary data $\mathbf{g}=\mathbf{0}$ and convergent approximations of the material parameters $M_h\longrightarrow M$, $M_h^{-1}\longrightarrow M^{-1}$ in $\mathrm{L}_\infty(\Omega;\mathbb{R}^{m\times m}_{sym})$, the discrete solutions $\big(\mathbf{u}_h\big)_{h\in\mathcal{H}}$ are converging to a weak solution $\mathbf{u}\in\mathrm{L}_2(Q;\mathbb{R}^m)$ of (2). Moreover, $\mathbf{u}$ is a strong solution satisfying (3), and the strong solution is unique.*

Proof. By the assumption $p_{n,K}=p_n$ we can apply Lem. 10 with the construction of the interpolation $\mathcal{I}_h$ and Cor. 12, so that $(\mathbf{u}_h)_{h\in\mathcal{H}}$ is uniformly bounded by

$$\big\|M_h^{1/2}\mathbf{u}_h\big\|_Q^2+T\,\big\|M_h^{1/2}\mathbf{u}_h(0)\big\|_\Omega^2\le4T\left(\big\|M_h^{-1/2}\mathbf{f}\big\|_Q^2+\big\|M_h^{1/2}\mathbf{u}_0\big\|_\Omega^2\right).$$

By (31) and the definition of ℓ_h (with $\mathbf{g}=\mathbf{0}$), this also implies that

$$\begin{aligned}
c_1\sum_{k=1}^m\big\|(\underline{A}_\mathbf{n}\mathbf{u}_h)_k\big\|^2_{(0,T)\times\Gamma_k}&=c_1\big\|\underline{A}_\mathbf{n}[\mathbf{u}_h]\big\|^2_{(0,T)\times\partial\Omega_h}\le b(\mathbf{u}_h,\mathbf{u}_h)=\langle\ell_h,\mathbf{u}_h\rangle=\big(\mathbf{f},\mathbf{u}_h\big)_Q+\big(M_h\mathbf{u}_0,\mathbf{u}_h(0)\big)_\Omega\\
&\le\frac12\big\|M_h^{-1/2}\mathbf{f}\big\|_Q^2+\frac12\big\|M_h^{1/2}\mathbf{u}_h\big\|_Q^2+\frac{1}{2T}\big\|M_h^{1/2}\mathbf{u}_0\big\|_\Omega^2+\frac{T}{2}\big\|M_h^{1/2}\mathbf{u}_h(0)\big\|_\Omega^2\\
&\le\Big(\frac12+2T\Big)\big\|M_h^{-1/2}\mathbf{f}\big\|_Q^2+\Big(\frac{1}{2T}+2T\Big)\big\|M_h^{1/2}\mathbf{u}_0\big\|_\Omega^2
\end{aligned}$$

is uniformly bounded for $h\in\mathcal{H}$, so that together with the asymptotic consistency of the material parameters $M_h\longrightarrow M$, $M_h^{-1}\longrightarrow M^{-1}$ in $\mathrm{L}_\infty(\Omega;\mathbb{R}^{m\times m}_{\mathrm{sym}})$ we obtain with a constant $C_{\mathbf{f},\mathbf{u}_0}>0$ depending on the data

$$\big\|M^{1/2}\mathbf{u}_h\big\|_Q^2+T\,\big\|M^{1/2}\mathbf{u}_h(0)\big\|_\Omega^2+c_1\sum_{k=1}^m\big\|(\underline{A}_\mathbf{n}\mathbf{u}_h)_k\big\|^2_{(0,T)\times\Gamma_k}\le C_{\mathbf{f},\mathbf{u}_0}\,,\qquad h\in\mathcal{H}\,.$$

The uniform stability in $\mathrm{L}_2(Q;\mathbb{R}^m)$ implies, that a subsequence $\mathcal{H}_0\subset\mathcal{H}$ with $0\in\overline{\mathcal{H}}_0$ and a weak limit $\mathbf{u}\in\mathrm{L}_2(Q;\mathbb{R}^m)$ with $\mathbf{u}(0)\in\mathrm{L}_2(\Omega;\mathbb{R}^m)$ and $(\underline{A}_\mathbf{n}\mathbf{u})_k|_{(0,T)\times\Gamma_k}\in\mathrm{L}_2((0,T)\times\Gamma_k)$ for $k=1,\dots,m$ exists, i.e.,

$$\begin{aligned}
\big(M\mathbf{u},\mathbf{v}\big)_Q&=\lim_{h\in\mathcal{H}_0}\big(M_h\mathbf{u}_h,\mathbf{v}\big)_Q\,, &&\mathbf{v}\in\mathrm{L}_2(Q;\mathbb{R}^m)\\
\big(M\mathbf{u}(0),\mathbf{v}_0\big)_\Omega&=\lim_{h\in\mathcal{H}_0}\big(M_h\mathbf{u}_h(0),\mathbf{v}_0\big)_\Omega\,, &&\mathbf{v}_0\in\mathrm{L}_2(\Omega;\mathbb{R}^m)\\
\big((\underline{A}_\mathbf{n}\mathbf{u})_k,v\big)_{(0,T)\times\Gamma_k}&=\lim_{h\in\mathcal{H}_0}\big((\underline{A}_\mathbf{n}\mathbf{u}_h)_k,v\big)_{(0,T)\times\Gamma_k}\,, &&v\in\mathrm{L}_2((0,T)\times\Gamma_k),\quad k=1,\dots,m\,.
\end{aligned}$$

Then we obtain

$$\big(\mathbf{u},L^*\mathbf{v}\big)_Q=\lim_{h\in\mathcal{H}_0}\big(\mathbf{u}_h,L^*\mathbf{v}\big)_{Q_h}=\lim_{h\in\mathcal{H}_0}\big(\mathbf{u}_h,L_h^*\mathbf{v}\big)_{Q_h}=\lim_{h\in\mathcal{H}_0}b_h(\mathbf{u}_h,\mathbf{v})\,,\qquad\mathbf{v}\in\mathcal{V}_h\,,$$

using dual consistency (30) for the last step. This extends to $\mathrm{H}^1_0(Q;\mathbb{R}^m)$, and by the assumption $p_{n,K},q_{n,K}\ge1$, for all $\mathbf{v}\in\mathrm{H}^1_0(Q;\mathbb{R}^m)$ a sequence $(\mathbf{v}_h)_{h\in\mathcal{H}_0}$ exists with $\mathbf{v}_h\in V_h\cap\mathrm{H}^1_0(Q;\mathbb{R}^m)$ and $\lim\limits_{h\in\mathcal{H}_0}\mathbf{v}_h=\mathbf{v}$, so that by (30)

$$\big(\mathbf{u},L^*\mathbf{v}\big)_Q=\lim_{h\in\mathcal{H}_0}b_h(\mathbf{u}_h,\mathbf{v})=\lim_{h\in\mathcal{H}_0}b_h(\mathbf{u}_h,\mathbf{v}_h)=\lim_{h\in\mathcal{H}_0}\big(\mathbf{f},\mathbf{v}_h\big)_Q=\big(\mathbf{f},\mathbf{v}\big)_Q\,,$$

i.e., for the limit $\mathbf{u}$ the weak derivative $L\mathbf{u} = \mathbf{f}$ in $\mathrm{L}_2(Q;\mathbb{R}^m)$ exists. This extends to initial and boundary data. Therefore, let $\overline{\mathcal{V}}^* \subset \mathrm{H}^1(Q;\mathbb{R}^m)$ be the closure of $\mathcal{V}^*$ in $\mathrm{H}^1(Q;\mathbb{R}^m)$; then, for all $\mathbf{v} \in \mathcal{V}^*$ a sequence $(\mathbf{v}_h)_{h\in\mathcal{H}_0}$ with $\mathbf{v}_h \in V_h \cap \overline{\mathcal{V}}^*$ and $\lim\limits_{h\in\mathcal{H}_0} \mathbf{v}_h = \mathbf{v}$ exists, and we get again by (30)

$$\big(\mathbf{u}, L^*\mathbf{v}\big)_Q = \lim_{h\in\mathcal{H}_0} b_h(\mathbf{u}_h,\mathbf{v}) = \lim_{h\in\mathcal{H}_0} b_h(\mathbf{u}_h,\mathbf{v}_h) = \lim_{h\in\mathcal{H}_0} \big\langle \ell_h, \mathbf{v}_h\big\rangle = \big(\mathbf{f},\mathbf{v}\big)_Q + \big(M\mathbf{u}_0, \mathbf{v}(0)\big)_\Omega .$$

Thus, using $\mathbf{v}(T) = \mathbf{0}$ for $\mathbf{v} = (v_1 \dots, v_m) \in \mathcal{V}^*$ yields

$$\begin{aligned} 0 &= \big(\mathbf{u}, L^*\mathbf{v}\big)_Q - \big(\mathbf{f},\mathbf{v}\big)_Q - \big(M\mathbf{u}_0,\mathbf{v}(0)\big)_\Omega = \big(\mathbf{u}, L^*\mathbf{v}\big)_Q - \big(L\mathbf{u},\mathbf{v}\big)_Q - \big(M\mathbf{u}_0,\mathbf{v}(0)\big)_\Omega \\ &= \big(M\mathbf{u}(0),\mathbf{v}(0)\big)_\Omega - \big(\underline{A}_\mathbf{n}\mathbf{u},\mathbf{v}\big)_{(0,T)\times\partial\Omega} - \big(M\mathbf{u}_0,\mathbf{v}(0)\big)_\Omega = \big(M(\mathbf{u}(0)-\mathbf{u}_0),\mathbf{v}(0)\big)_\Omega + \sum_{k=1}^m \big((\underline{A}_\mathbf{n}\mathbf{u})_k, v_k\big)_{(0,T)\times\Gamma_k} , \end{aligned}$$

so that $\mathbf{u}(0) = \mathbf{u}_0$ in Ω and $(\underline{A}_\mathbf{n}\mathbf{u})_k = 0$ on $(0,T)\times\Gamma_k$ for $k = 1,\dots,m$, and thus $\mathbf{u}$ is indeed a strong solution with homogeneous boundary conditions at $(0,T)\times\partial\Omega$.

Next, we show that the weak limit is unique. Therefore, select another subsequence $\mathcal{H}_1 \subset \mathcal{H}$ with $0 \in \overline{\mathcal{H}}_1$ and with a weak limit $\tilde{\mathbf{u}} \in \mathrm{L}_2(Q;\mathbb{R}^m)$ with $\tilde{\mathbf{u}}(0) \in \mathrm{L}_2(\Omega;\mathbb{R}^m)$ and $(\underline{A}_\mathbf{n}\tilde{\mathbf{u}})_k|_{(0,T)\times\Gamma_k} \in \mathrm{L}_2((0,T)\times\Gamma_k)$ for $k = 1,\dots,m$. Then, we also obtain $\tilde{\mathbf{u}}(0) = \mathbf{u}_0$ and $(\underline{A}_\mathbf{n}\tilde{\mathbf{u}})_k = 0$ for $k = 1,\dots,m$. A sequence $(\mathbf{e}_h)_{h\in\mathcal{H}}$ with $\mathbf{e}_h \in V_h$ exists such that $\lim_{h\in\mathcal{H}} \mathbf{e}_h = \mathbf{u}-\tilde{\mathbf{u}}$, and we get

$$\begin{aligned} \frac{1}{2}\big\|M^{1/2}(\mathbf{u}-\tilde{\mathbf{u}})\big\|_Q^2 = \frac{1}{2}\lim_{h\in\mathcal{H}} \big\|M^{1/2}\mathbf{e}_h\big\|_Q^2 &\le \lim_{h\in\mathcal{H}} b_h\big(\mathbf{e}_h, \mathcal{I}_h(d_T\mathbf{e}_h)\big) = \lim_{h\in\mathcal{H}_0} b_h\big(\mathbf{u}_h, \mathcal{I}_h(d_T\mathbf{e}_h)\big) - \lim_{h\in\mathcal{H}_1} b_h\big(\tilde{\mathbf{u}}, \mathcal{I}_h(d_T\mathbf{e}_h)\big) \\ &= \lim_{h\in\mathcal{H}_0} \big\langle \ell_h, \mathcal{I}_h(d_T\mathbf{e}_h)\big\rangle - \lim_{h\in\mathcal{H}_1} \big\langle \ell_h, \mathcal{I}_h(d_T\mathbf{e}_h)\big\rangle \\ &= \big\langle \ell, d_T(\mathbf{u}-\tilde{\mathbf{u}})\big\rangle - \big\langle \ell, d_T(\mathbf{u}-\tilde{\mathbf{u}})\big\rangle = 0 , \end{aligned}$$

so that $\mathbf{u} = \tilde{\mathbf{u}}$. This shows that the weak limit is unique, so that the full sequence is converging, i.e., $\lim_{h\in\mathcal{H}} \mathbf{u}_h = \mathbf{u}$. The same argument applies to all strong solutions, i.e., $\mathbf{u}$ is the unique strong solution of (3). □

Remark 16. *The result extends to inhomogeneous boundary data* $\mathbf{g} \neq \mathbf{0}$, *if an extension* $\mathbf{u}_\mathbf{g} \in \mathrm{L}_2(Q;\mathbb{R}^m)$ *exists with* $L\mathbf{u}_\mathbf{g} \in \mathrm{L}_2(Q;\mathbb{R}^m)$ *and* $(A_\mathbf{n}\mathbf{u}_\mathbf{g})_k \in \mathrm{L}_2(I\times\Gamma_k)$ *satisfying* $(A_\mathbf{n}\mathbf{u}_\mathbf{g})_k = g_k$, $k = 1,\dots,m$. *In particular, the regularity result that the limit of the DG approximations is a strong solution requires sufficient regularity of the boundary data.*

5.2. Convergence in the DG norm

We adapt the convergence result for the DG norm (40) in [Di Pietro and Ern, 2011, Thm. 2.37] to our setting.

Theorem 17. *Assume that the strong solution of* (3) *is sufficiently smooth satisfying* $\mathbf{u} \in \mathrm{H}^s(Q;\mathbb{R}^m)$ *with* $s \geq 1$. *Then, the error for the discrete solution* $\mathbf{u}_h \in V_h$ *of* (35) *is bounded by*

$$\|\mathbf{u}-\mathbf{u}_h\|_{h,\mathrm{DG}} \leq C h^{s-1/2}\,\|\mathrm{D}^s\mathbf{u}\|_Q + CTh^{-1/2}\|M_h^{-1/2}(M_h - M)\partial_t\mathbf{u}\|_Q$$

with $C > 0$ *depending on the mesh regularity, the polynomial degree, and the material parameters.*

Proof. Since we assume for the solution $\mathbf{u} \in \mathrm{H}^1(Q;\mathbb{R}^m)$, we have $L\mathbf{u}, L_h\mathbf{u} \in \mathrm{L}_2(Q;\mathbb{R}^m)$, for all traces $\mathbf{u}|_{\partial Q_h} \in \mathrm{L}_2(\partial Q_h;\mathbb{R}^m)$, $[\mathbf{u}]_n = \mathbf{0}$ for $n = 1,\dots,N-1$, and $A_{\mathbf{n}}[\mathbf{v}] = \mathbf{0}$ on $I_h \times F$ for inner faces $F \in \mathcal{F}_h \setminus \partial\Omega$, and $(\underline{A}_{\mathbf{n}}\mathbf{u})_k = g_k$ on $I \times \Gamma_k$ for $k = 1,\dots,m$, so that $b_h(\mathbf{u},\mathbf{v}_h)$ is well defined with

$$b_h(\mathbf{u},\mathbf{w}_h) = \big(L_h\mathbf{u},\mathbf{w}_h\big)_Q + (M_h\mathbf{u}(0),\mathbf{w}_h)_Q + \int_0^T \langle \ell_{\partial\Omega,h}(t),\mathbf{w}_h\rangle\,\mathrm{d}t = \langle \ell_h,\mathbf{w}_h\rangle + \big((M_h - M)\partial_t\mathbf{u},\mathbf{w}_h\big)_Q\,, \qquad \mathbf{w}_h \in V_h\,. \quad (46)$$

Thus we obtain for the discrete solution $\mathbf{u}_h \in V_h$ Galerkin orthogonality up to data error

$$b_h(\mathbf{u}_h,\mathbf{w}_h) = b_h(\mathbf{u},\mathbf{w}_h) + \big((M - M_h)\partial_t\mathbf{u},\mathbf{w}_h\big)_Q\,, \qquad \mathbf{w}_h \in V_h\,.$$

By the trace estimate (12) we obtain $\|\mathbf{w}_h\|_{h,\mathrm{DG}^+}^2 \leq (C_{\mathrm{tr}}^2+1)h^{-1}\|M_h^{1/2}\mathbf{w}_h\|_Q^2$, so that by Lem. 7

$$\begin{aligned}\|M_h^{1/2}\mathbf{w}_h\|_Q^2 &\leq 2\,b_h(\mathbf{w}_h,d_T\mathbf{w}_h) \leq 2\,\|\mathbf{w}_h\|_{h,\mathrm{DG}}\|d_T\mathbf{w}_h\|_{h,\mathrm{DG}^+} \leq 2T\,\|\mathbf{w}_h\|_{h,\mathrm{DG}}\|\mathbf{w}_h\|_{h,\mathrm{DG}^+}\\ &\leq 2T^2(C_{\mathrm{tr}}^2+1)h^{-1}\|\mathbf{w}_h\|_{h,\mathrm{DG}}^2 + \frac{1}{2(C_{\mathrm{tr}}^2+1)}h\,\|\mathbf{w}_h\|_{h,\mathrm{DG}^+}^2\\ &\leq 2T^2(C_{\mathrm{tr}}^2+1)h^{-1}\,\|\mathbf{w}_h\|_{h,\mathrm{DG}}^2 + \frac{1}{2}\|M_h^{1/2}\mathbf{w}_h\|_Q^2\,,\end{aligned}$$

so that the consistency term can by bounded by

$$\begin{aligned}\big((M - M_h)\partial_t\mathbf{u},\mathbf{w}_h\big)_Q &\leq \|\big(M_h^{-1/2}(M_h - M)\partial_t\mathbf{u}\|_Q\|M_h^{1/2}\mathbf{w}_h\|_Q\\ &\leq 2T\sqrt{C_{\mathrm{tr}}^2+1}h^{-1/2}\|\big(M_h^{-1/2}(M_h - M)\partial_t\mathbf{u}\|_Q\|\mathbf{w}_h\|_{h,\mathrm{DG}}\,.\end{aligned}$$

For all $\mathbf{v}_h \in V_h$ this yields the estimate, using Thm. 14 and continuity of the bilinear form $b_h(\cdot,\cdot)$ in the DG norms

$$\begin{aligned}c_{\text{inf-sup}}\|\mathbf{u}_h - \mathbf{v}_h\|_{h,\mathrm{DG}} &\leq \sup_{\mathbf{w}_h\in V\setminus\{\mathbf{0}\}} \frac{b_h(\mathbf{u}_h - \mathbf{v}_h,\mathbf{w}_h)}{\|\mathbf{w}_h\|_{h,\mathrm{DG}}} = \sup_{\mathbf{w}_h\in V\setminus\{\mathbf{0}\}} \frac{b_h(\mathbf{u}-\mathbf{v}_h,\mathbf{w}_h) + \big((M - M_h)\partial_t\mathbf{u},\mathbf{w}_h\big)_Q}{\|\mathbf{w}_h\|_{h,\mathrm{DG}}}\\ &\leq \|\mathbf{u}-\mathbf{v}_h\|_{h,\mathrm{DG}^+} + 2T\sqrt{C_{\mathrm{tr}}^2+1}h^{-1/2}\|\big(M_h^{-1/2}(M_h - M)\partial_t\mathbf{u}\|_Q\,.\end{aligned}$$

Now select an H^1-stable quasi-interpolation $\mathbf{v}_h = \Pi_h^{\mathrm{Cl}}\mathbf{u}$ of Clement-type [Bartels, 2016, Sect. 4.4.2] with

$$\|M^{1/2}(\mathbf{u}-\Pi_h^{\mathrm{Cl}}\mathbf{u})\|_Q \leq C_4 h\,\|\mathrm{D}\mathbf{u}\|_Q\,, \qquad \|M^{-1/2}L_h(\mathbf{u}-\Pi_h^{\mathrm{Cl}}\mathbf{u})\|_Q \leq C_5\|\mathrm{D}\mathbf{u}\|_Q$$

and

$$\|M^{1/2}(\mathbf{u}-\Pi_h^{\mathrm{Cl}}\mathbf{u})\|_{\partial Q_h} + h^{-1/2}\|M^{1/2}(\mathbf{u}-\Pi_h^{\mathrm{Cl}}\mathbf{u})\|_Q + h^{1/2}\|M^{-1/2}L_h(\mathbf{u}-\Pi_h^{\mathrm{Cl}}\mathbf{u})\|_Q \leq C_6 h^{s-1/2}\|\mathrm{D}^s\mathbf{u}\|_Q\,.$$

Then, the result follows from interpolation estimates using [Di Pietro and Ern, 2011, Lem. 1.59] and

$$\begin{aligned}\|\mathbf{u}-\mathbf{u}_h\|_{h,\mathrm{DG}} &\leq \|\mathbf{u}-\Pi_h^{\mathrm{Cl}}\mathbf{u}\|_{h,\mathrm{DG}} + \|\mathbf{u}_h - \Pi_h^{\mathrm{Cl}}\mathbf{u}\|_{h,\mathrm{DG}}\\ &\leq \|\mathbf{u}-\Pi_h^{\mathrm{Cl}}\mathbf{u}\|_{h,\mathrm{DG}} + c_{\text{inf-sup}}^{-1}\|\mathbf{u}-\Pi_h^{\mathrm{Cl}}\mathbf{u}\|_{h,\mathrm{DG}^+} + 2T\sqrt{C_{\mathrm{tr}}^2+1}\,c_{\text{inf-sup}}^{-1}h^{-1/2}\|M_h^{-1/2}(M_h - M)\partial_t\mathbf{u}\|_Q\\ &\leq C_6 h^{s-1/2}\|\mathrm{D}^s\mathbf{u}\|_Q + C_7 T h^{-1/2}\|M_h^{-1/2}(M_h - M)\partial_t\mathbf{u}\|_Q\,.\end{aligned}$$

□

This recovers the convergence result [Bansal et al., 2021, Prop. 6.5] for the DG semi-norm (40).

Corollary 18. *Assume that the strong solution of (3) is sufficiently smooth satisfying $\mathbf{u} \in \mathrm{H}^s(Q;\mathbb{R}^m)$ with $s \geq 1$. Then, the error for the discrete solution $\mathbf{u}_h \in V_h$ of (35) is bounded in every time step by*

$$\big\|M_h^{1/2}\big(\mathbf{u}(t_n) - \mathbf{u}_h(t_n)\big)\big\|_\Omega \leq Ch^{s-1/2}\,\big\|\mathrm{D}^s\mathbf{u}\big\|_{(0,t_n)\times\Omega} + CTh^{-1/2}\big\|M_h^{-1/2}(M_h - M)\partial_t\mathbf{u}\big\|_{(0,t_n)\times\Omega}$$

with $C > 0$ depending on the mesh regularity, the polynomial degree, and the material parameters.

For the proof Thm. 17 is applied with $T = t_n$; then, the assertion directly follows from $\frac{1}{2}\big\|M_h^{1/2}\mathbf{v}(T)\big\|_\Omega \leq \big\|\mathbf{v}\big\|_{h,\mathrm{DG}}$.

Remark 19. *If $M \in \mathrm{L}_\infty(\Omega;\mathbb{R}^{m\times m}_{\mathrm{sym}})$ is smooth, the consistency term can be estimated by*

$$\|\big(M_h^{-1/2}(M_h - M)\partial_t\mathbf{u}\big\|_Q \leq \big\|M_h^{-1/2}(M - M_h)M^{-1/2}\big\|_\infty\big\|M^{1/2}\partial_t\mathbf{u}\big\|_Q\,.$$

If M is discontinuous and if the jumps of the material parameters are not resolved by the mesh, the consistency error can be estimated in case of higher regularity of the solution: if $\partial_t\mathbf{u} \in \mathrm{L}_2(0,T;\mathrm{L}_q(\Omega;\mathbb{R}^m))$ with $q > 2$, we obtain

$$\big\|M_h^{-1/2}(M_h - M)\partial_t\mathbf{u}\big\|_Q \leq \big\|M_h^{-1/2}(M - M_h)M^{-1/2}\big\|_{\mathrm{L}_{2q/(2-q)}(\Omega;\mathbb{R}^{m\times m}_{\mathrm{sym}})}\big\|M^{1/2}\partial_t\mathbf{u}\big\|_{\mathrm{L}_2(0,T;\mathrm{L}_q(\Omega;\mathbb{R}^m))}\,. \tag{47}$$

5.3. Error control

For the error $\mathbf{u} - \mathbf{u}_h$ in the DG semi-norm we obtain from (18) and (20)

$$\begin{aligned}\big|\mathbf{u} - \mathbf{u}_h\big|^2_{h,\mathrm{DG}} \leq \frac{1}{2}\Bigg(&\big\|M_h^{1/2}(\mathbf{u}_h(0) - \mathbf{u}_0)\big\|^2_\Omega + \sum_{n=1}^{N-1}\big\|M_h^{1/2}[\mathbf{u}_h]_n\big\|^2_\Omega + \big\|M_h^{1/2}(\mathbf{u}_h(T) - \mathbf{u}(T))\big\|^2_\Omega\Bigg)\\ &+ \sum_{k=1}^m\big\|(A_\mathbf{n}\mathbf{u}_h)_k - g_k\big\|^2_{I_h\times\Gamma_k} + C_1\big\|A_\mathbf{n}[\mathbf{u}_h]\big\|^2_{I_h\times(\partial\Omega_h\cap\Omega)}\end{aligned} \tag{48}$$

and in the DG norm

$$\begin{aligned}\big\|\mathbf{u} - \mathbf{u}_h\big\|^2_{h,\mathrm{DG}} &= \big|\mathbf{u} - \mathbf{u}_h\big|^2_{h,\mathrm{DG}} + \big\|h^{1/2}M_h^{-1/2}L_h(\mathbf{u} - \mathbf{u}_h)\big\|^2_{Q_h}\\ &\leq \big|\mathbf{u} - \mathbf{u}_h\big|^2_{h,\mathrm{DG}} + 2\,\big\|h^{1/2}M_h^{-1/2}(L_h\mathbf{u}_h - \mathbf{f})\big\|^2_{Q_h} + 2\,\big\|h^{1/2}M_h^{-1/2}(M - M_h)\partial_t\mathbf{u}\big\|^2_{Q_h}\,.\end{aligned} \tag{49}$$

Up to the error $\mathbf{u}_h - \mathbf{u}$ at final time T in (48) and the parameter approximation error $M - M_h$ in (49), this can be evaluated explicitly by the residual error indicator $\eta_{\mathrm{res},h} = \Big(\sum_{R\in\mathcal{R}_h}\eta^2_{\mathrm{res},R}\Big)^{1/2}$ given by the local contributions

$$\eta^2_{\mathrm{res},R} = \eta^2_{\mathrm{res},n,K} + 2h_K\,\big\|M_h^{-1/2}(L_h\mathbf{u}_h - \mathbf{f})\big\|^2_R + \sum_{k=1}^m\big\|(A_\mathbf{n}\mathbf{u}_h)_k - g_k\big\|^2_{(t_{n-1},t_n)\times(\Gamma_k\cap\partial K)} + C_1\big\|A_\mathbf{n}[\mathbf{u}_h]\big\|^2_{(t_{n-1},t_n)\times(\Omega\cap\partial K)}$$

for $R = (t_{n-1},t_n)\times K$, $n = 1,\dots,N$, with

$$\begin{aligned}\eta^2_{\mathrm{res},1,K} &= \frac{1}{2}\big\|M_h^{1/2}(\mathbf{u}_h(0) - \mathbf{u}_0)\big\|^2_K + \frac{1}{2}\big\|M_h^{1/2}[\mathbf{u}_h]_1\big\|^2_K\,, && R = (0,t_1)\times K\,,\\ \eta^2_{\mathrm{res},n,K} &= \frac{1}{2}\big\|M_h^{1/2}[\mathbf{u}_h]_{n-1}\big\|^2_K + \frac{1}{2}\big\|M_h^{1/2}[\mathbf{u}_h]_n\big\|^2_K\,, && R = (t_{n-1},t_n)\times K\,,\ 1 < n < N\,,\\ \eta^2_{\mathrm{res},N,K} &= \frac{1}{2}\big\|M_h^{1/2}[\mathbf{u}_h]_{N-1}\big\|^2_K\,, && R = (t_{N-1},T)\times K\,.\end{aligned}$$

Lemma 20. *Let $\mathbf{u} \in \mathrm{L}_2(Q;\mathbb{R}^m)$ be the weak solution of (2) and $\mathbf{u}_h \in V_h$ the discrete solution of (35). Then, if $\mathbf{u}$ is a strong solution, the error in the DG norm is bounded by*

$$\big\|\mathbf{u} - \mathbf{u}_h\big\|_{h,\mathrm{DG}} \leq \Big(\eta^2_{\mathrm{res},h} + \big\|M_h^{1/2}(\mathbf{u}_h(T) - \mathbf{u}(T))\big\|^2_\Omega + 2\,\big\|h^{1/2}M_h^{-1/2}(M - M_h)\partial_t\mathbf{u}\big\|^2_{Q_h}\Big)^{1/2}.$$

6. Numerical experiments

The convergence estimates are illustrated by numerical experiments for acoustics (5) with Dirichlet boundary conditions.
Experiment 1. We test the convergence of the solution with smooth initial value and piecewise constant material

$$\varrho(\mathbf{x}) = \begin{cases} 1 & \mathbf{x}\cdot\mathbf{m} \le \gamma\,, \\ 2 & \mathbf{x}\cdot\mathbf{m} > \gamma\,, \end{cases} \quad \kappa(\mathbf{x}) = 1/\varrho(\mathbf{x})\,, \quad \mathbf{f} = \mathbf{0}\,, \quad Q = (0,1)\times(0,1)^2\,, \quad \gamma \in (0,1)\,, \qquad \mathbf{m}\in\mathbb{R}^2\,,\ \mathbf{m}\cdot\mathbf{m} = 1\,,$$

so that the impedance is constant across the interface. We start with

$$\mathbf{u}_0(\mathbf{x}) = a_0(\mathbf{x}\cdot\mathbf{m})\begin{pmatrix}1\\ \mathbf{m}\end{pmatrix} \qquad \text{with} \qquad a_0(x) = \begin{cases}\sin(3\pi x)^2 & x\in[0,1/3]\\ 0 & \text{else.}\end{cases}$$

Then, the solution is given by $\mathbf{u}(t,\mathbf{x}) = \begin{cases}\mathbf{u}_0(\mathbf{x}-t\mathbf{m}) & \mathbf{x}\cdot\mathbf{m}\le\gamma\,,\\ \mathbf{u}_0(2\mathbf{x}-(t+2/3)\mathbf{m}) & \mathbf{x}\cdot\mathbf{m}>\gamma\,.\end{cases}$

Case a). If the material interface is resolved by the mesh ($M = M_h$), we observe for linear approximations in space and time on uniformly refined meshes the expected convergence rate in the DG norm (Fig. 1). For this configuration also the dual problem is smooth which results in better convergence rates for the L_2 error, in particular in the adaptive case.

ℓ	$\|\mathbf{u}-\mathbf{u}_h\|_{\mathrm{DG},h}$	$\|M_h^{1/2}(\mathbf{u}-\mathbf{u}_h)\|_Q$
3	0.744239	0.244213
4	0.569868	0.135738
5	0.328281	0.047846
6	0.143545	0.011040
7	0.054274	0.002494
8	0.019645	0.000615

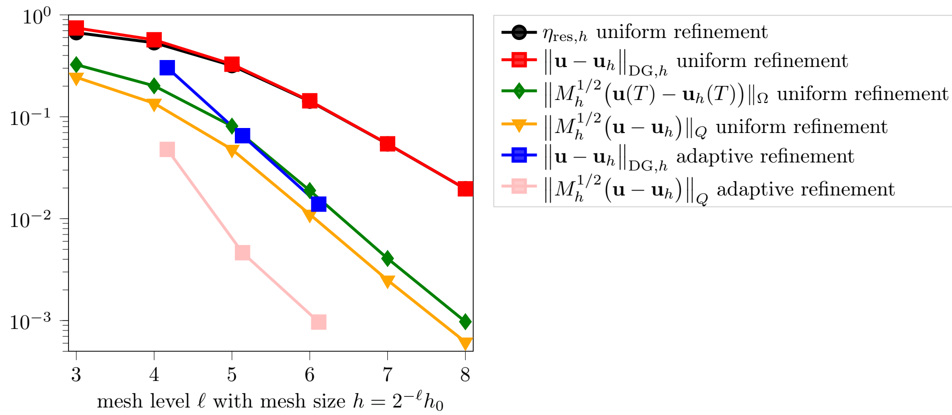


Figure 1. Convergence test for the first experiment with $\gamma = 0.5$ and $\mathbf{m} = \begin{pmatrix}0\\1\end{pmatrix}$.
.

Case b). If the material interface cannot be resolved by the mesh ($M \ne M_h$), the consistency error gets relevant, which is observed by the results in Fig. 2.

ℓ	$\|\mathbf{u}-\mathbf{u}_h\|_{\mathrm{DG},h}$	$\|M_h^{1/2}(\mathbf{u}-\mathbf{u}_h)\|_Q$
3	0.803100	0.222789
4	0.566667	0.111046
5	0.298943	0.035623
6	0.126032	0.012112
7	0.051102	0.006775
8	0.022482	0.004264

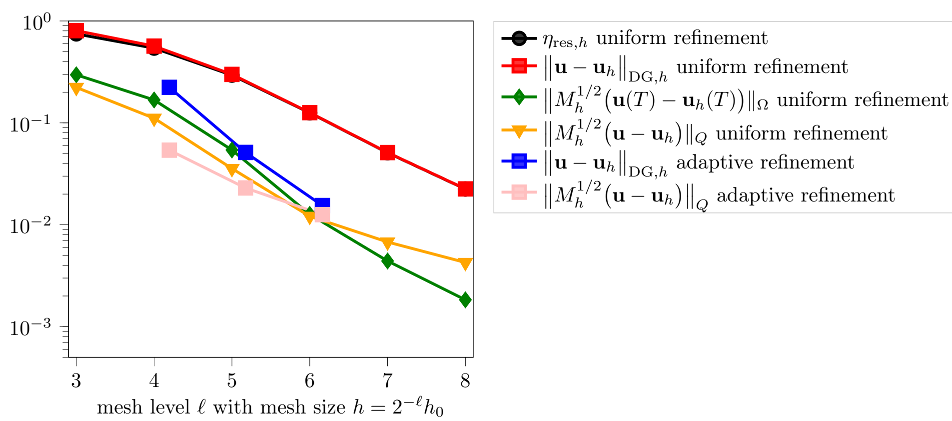


Figure 2. Convergence test for the first experiment with $\gamma = 4/7$ and $\mathbf{m} = \begin{pmatrix}0.8\\0.6\end{pmatrix}$.
.

Although the material interface cannot be resolved by the mesh, the solution is sufficiently smooth so that the approximation error of the material data $M_h - M$ can be estimated by Rem. 19. We observe nearly optimal convergence in the DG norm, but now the L_2 convergence gets worse in comparison with the first case.

In both cases, the convergence of $\mathbf{u}(T) - \mathbf{u}_h(T)$ in L_2 is faster than the convergence in the DG norm, and the residual error indicator yield results close to the error in the DG norm; this confirms the estimate in Lem. 20 . We observe that adaptivity provides better solutions with a substantial reduction of the required problem size $\dim V_h$ to achieve a certain accuracy. Therefore a single adaptive step is sufficient, where the polynomial degree in space and time is increased for $\eta_{\mathrm{res},R} \geq \vartheta_1 \max_{R' \in \mathrm{R}_h} \eta_{\mathrm{res},R'}$ and decreased for $\eta_{\mathrm{res},R} \leq \vartheta_0 \max_{R' \in \mathrm{R}_h} \eta_{\mathrm{res},R'}$, depending on $\vartheta_1 > \vartheta_0 > 0$. Note that this results in a different refinement pattern in every time interval, and a simple refinement in space is not sufficient for a strong reduction of the required degrees of unknowns. Here, we select $\vartheta_1 = 0.3$ and $\vartheta_0 = 0.02$, and in the figures for the adaptive results the mesh size is logarithmically interpolated depending on the degrees of freedom.

Experiment 2. Finally we test the convergence of a Riemann problem, where the solution is given by

$$\mathbf{u}(t,\mathbf{x}) = \begin{cases} \begin{pmatrix} 0 \\ \mathbf{0} \end{pmatrix} & \mathbf{x}\cdot\mathbf{m} < -t\,, \\ \begin{pmatrix} 1 \\ \mathbf{m} \end{pmatrix} & -t < \mathbf{x}\cdot\mathbf{m} < t\,, \\ \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix} & t < \mathbf{x}\cdot\mathbf{m}\,, \end{cases} \qquad \mathbf{m} = \begin{pmatrix} 0.8 \\ 0.6 \end{pmatrix}, \quad \kappa = 1\,, \quad \varrho = 1\,, \quad \mathbf{f} = \mathbf{0}\,, \quad Q = (0,1/2)\times(-1,1)\times(0,1)\,.$$

Then, $L\mathbf{u} = \mathbf{0}$, so that $\mathbf{u}$ is a strong solution, and since the condition in Rem. 16 applies, we obtain convergence in the limit by Thm. 15. On the other hand, the solution is piecewise discontinuous, so that the smoothness assumption in Thm. 17 is not satisfied. We also observe convergence, cf. Fig. 3, but with a reduced rate $\mathcal{O}(h^{1/3})$. In particular, the rate is not improved for the L_2 error, and simple adaptivity is not sufficient to increase the efficiency.

ℓ	$\|\mathbf{u} - \mathbf{u}_h\|_{\mathrm{DG},h}$	$\|M_h^{1/2}(\mathbf{u} - \mathbf{u}_h)\|_Q$
3	0.965539	0.201996
4	0.769326	0.152608
5	0.605518	0.116801
6	0.474996	0.089880
7	0.371885	0.069248
8	0.290613	0.053403

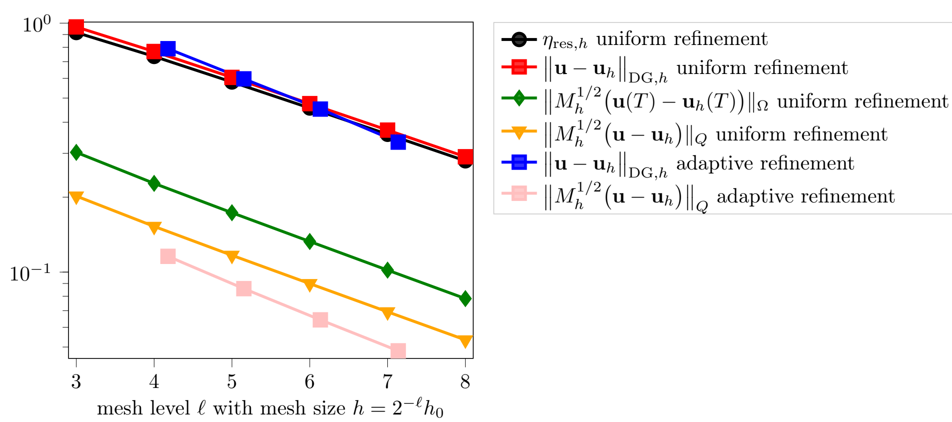


Figure 3. Convergence test for the Riemann Problem.

.

Here, the solution is not smooth, and the results do not improve if the material parameters are aligned with the mesh. Moreover, further tests show that the convergence rate $\mathcal{O}(h^{1/3})$ in the DG norm cannot be improved by adaptivity, which indicates that without sufficient regularity and jumps along the characteristics the DG norm is not appropriate for a qualitative convergence analysis, as it is possible for point singularities, see [Bansal et al., 2021]. Then, the convergence analysis requires high regularity in weighted Sobolev spaces.

7. Conclusion and Outlook

The convergence analysis in the DG norm only assumes regularity of the space-time solution $\mathbf{u}$ in $\mathrm{H}^1(Q;\mathbb{R}^m)$; this implies regularity of the solution $\mathbf{u}(t_n)$ at all time steps in $\mathrm{H}^{1/2}(\Omega;\mathbb{R}^m)$. This clearly extends convergence results with respect to the graph norm, where the analysis requires higher regularity. Moreover, the simple residual error indicator yields estimates very close to the error in the DG norm. On the other hand, for discontinuous Riemann problems we can prove only convergence in the limit, and the numerical experiments demonstrate that we obtain convergence in L_2 but with a reduced rate, which can be improved by adaptivity in L_2 but not in the DG norm.

All our estimates rely on a Hilbert space setting. This may be not appropriate for hyperbolic systems, and numerical tests demonstrate better convergence rates in $\mathrm{L}_1(Q;\mathbb{R}^m)$, but a corresponding analysis remains an open problem.

Acknowledgement

The authors gratefully acknowledge the support of the Deutsche Forschungsgemeinschaft (DFG) within the SFB 1173 "Wave Phenomena" (Project-ID 258734477).

Data availability statement

All data are available in https://git.scc.kit.edu/mpp/mpp/-/tags/st-experiments.

Conflict of interest

The authors declare no potential conflict of interests.

References

[Babuška et al., 2001] Babuška, I., Feistauer, M., and Šolín, P. (2001). On one approach to a posteriori error estimates for evolution problems solved by the method of lines. *Numerische Mathematik*, 89(2):225–256.

[Bansal et al., 2021] Bansal, P., Moiola, A., Perugia, I., and Schwab, C. (2021). Space–time discontinuous Galerkin approximation of acoustic waves with point singularities. *IMA Journal of Numerical Analysis*, 41(3):2056–2109.

[Bartels, 2016] Bartels, S. (2016). *Numerical Approximation of Partial Differential Equations*, volume 64. Springer.

[Baumgarten and Wieners, 2021] Baumgarten, N. and Wieners, C. (2021). The parallel finite element system M++ with integrated multilevel preconditioning and multilevel Monte Carlo methods. *Computers & Mathematics with Applications*, 81:391–406.

[Di Pietro and Ern, 2011] Di Pietro, D. A. and Ern, A. (2011). *Mathematical aspects of discontinuous Galerkin methods*, volume 69. Springer.

[Dörfler et al., 2016] Dörfler, W., Findeisen, S., and Wieners, C. (2016). Space-time discontinuous Galerkin discretizations for linear first-order hyperbolic evolution systems. *Comput. Methods Appl. Math.*, 16(3):409–428.

[Dörfler et al., 2019] Dörfler, W., Findeisen, S., Wieners, C., and Ziegler, D. (2019). Parallel adaptive discontinuous Galerkin discretizations in space and time for linear elastic and acoustic waves. In Langer, U. and Steinbach, O., editors, *Space-Time Methods. Applications to Partial Differential Equations*, volume 25 of *Radon Series on Computational and Applied Mathematics*, pages 97–127. Walter de Gruyter.

[Dörfler et al., 2020] Dörfler, W., Wieners, C., and Ziegler, D. (2020). Space-time discontinuous Galerkin methods for linear hyperbolic systems and the application to the forward problem in seismic imaging. In Klöfkorn, R., Keilegavlen, E., Radu, F., and Fuhrmann, J., editors, *Finite Volumes for Complex Applications IX – Methods, Theoretical Aspects, Examples*, volume 323 of *Springer Proceedings in Mathematics & Statistics*, pages 477–485. Springer.

[Ern and Guermond, 2021] Ern, A. and Guermond, J.-L. (2021). *Finite elements III: First-Order and Time-Dependent PDEs*, volume 74. Springer.

[Ernesti and Wieners, 2019a] Ernesti, J. and Wieners, C. (2019a). Space-time discontinuous Petrov–Galerkin methods for linear wave equations in heterogeneous media. *Computational Methods in Applied Mathematics*, 19(3):465–481.

[Ernesti and Wieners, 2019b] Ernesti, J. and Wieners, C. (2019b). A space-time DPG method for acoustic waves. In Langer, U. and Steinbach, O., editors, *Space-Time Methods. Applications to Partial Differential Equations*, volume 25 of *Radon Series on Computational and Applied Mathematics*, pages 99–127. Walter de Gruyter.

[Falk and Richter, 1999] Falk, R. S. and Richter, G. R. (1999). Explicit finite element methods for symmetric hyperbolic equations. *SIAM J. Numer. Anal.*, 36(3):935–952.

[Gopalakrishnan et al., 2017] Gopalakrishnan, J., Schöberl, J., and Wintersteiger, C. (2017). Mapped tent pitching schemes for hyperbolic systems. *SIAM J. Sci. Comput.*, 39(6):B1043–B1063.

[Hochbruck et al., 2015] Hochbruck, M., Pažur, T., Schulz, A., Thawinan, E., and Wieners, C. (2015). Efficient time integration for discontinuous Galerkin approximations of linear wave equations. *ZAMM*, 95(3):237–259.

[Jovanović and Rohde, 2005] Jovanović, V. and Rohde, C. (2005). Finite-volume schemes for Friedrichs systems in multiple space dimensions: A priori and a posteriori error estimates. *Numerical Methods for Partial Differential Equations: An International Journal*, 21(1):104–131.

[Rauch, 1985] Rauch, J. (1985). On convergence of the finite element method for the wave equation. *SIAM J. Numer. Anal.*, 22(2):245–249.

[Schafelner, 2022] Schafelner, A. (2022). *Space-Time Finite Element Methods*. PhD thesis, Johannes Kelper University Linz.