Karlsruhe Institute of Technology

Wave phenomena
analysis and numerics

A space-time Petrov-Galerkin method for linear wave equations

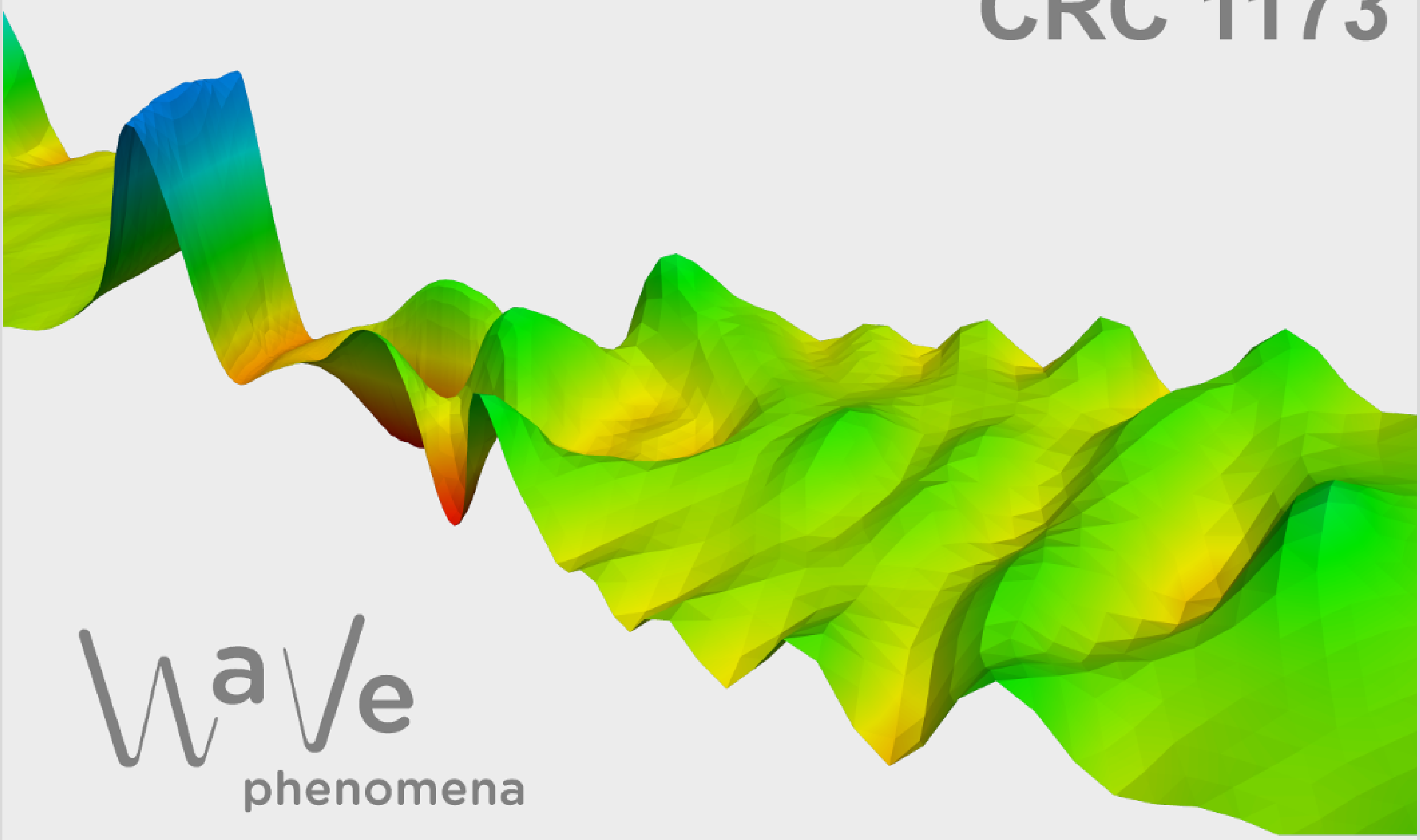
(Lecture notes for Zurich Summer School 2016 „Numerical Methods for Wave Propagation")

Christian Wieners

CRC Preprint 2018/19, August 2018

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173



Wave phenomena

Participating universities

Universität Stuttgart

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Funded by

DFG

ISSN 2365-662X

We introduce a space-time discretization for linear first-order hyperbolic evolution systems using a discontinuous Galerkin approximation in space and a Petrov–Galerkin scheme in time. For the dG method, the upwind flux is evaluated by explicitly solving a Riemann problem. Then we show well-posedness and convergence of the discrete system. Based on goal-oriented dual-weighted error estimation an adaptive strategy is introduced. The full space-time linear system is solved with a parallel multilevel preconditioner. Numerical experiments for acoustic and electro-magnetic waves underline the efficiency of the overall adaptive solution process.

These notes are an extended version of *Space-time discontinuous Galerkin discretizations for linear first-order hyperbolic evolution systems* by W. Dörfler, S. Findeisen, and C. Wieners [DFW16], combined with material from the overview on time integration given in *Efficient time integration for discontinuous Galerkin approximations of linear wave equations* by M. Hochbruck, T. Pazur, A. Schulz, E. Thawinan, and C. Wieners [HPS+15]. Basic results on discontinuous Galerkin methods are taken from the textbooks by R.J. Leveque [Lev02] and J.S. Hesthaven and T. Warburton [HW08].

Acknowledgments

We gratefully acknowledge the support of the German Research Foundation (DFG) by GRK 1294, FOR 1650, and CRC 1173.

Contents

1 Linear wave equations

We introduce several examples for wave equations. Starting with the simple homogeneous 1D case, we then consider elastic waves in solids and fluid, and electro-magnetic waves in vacuum. In all these cases, the physical setting results in first-order evolution systems. Here we only consider configurations and applications with small energy so that linearized constitutive equations describe the wave propagation sufficiently accurate.

A simple 1D wave The most simple equation describing a wave is given by

$$\partial_t^2 \varphi(t,x) = c^2 \partial_x^2 \varphi(t,x) \,,$$

where $t \in (0,T)$ is the time variable, $t = 0$ is the initial time, $T > 0$ is the final time, and $x \in \mathbb{R}$ is the position on the real line. The displacement at time t and position x is denoted by $\varphi(t,x)$, and $c > 0$ is the wave speed.

For given initial displacement $\varphi(0,\cdot)$ and velocity $\partial_t\varphi(0,\cdot)$ the solution is explicitly given by d'Alembert formula

$$\varphi(t,x) = \frac{1}{2}\left(\varphi(0,x-ct) + \varphi(0,x+ct) + \frac{1}{c}\int_{x-ct}^{x+ct} \partial_t\varphi(0,\xi)\,\mathrm{d}\xi\right) .$$

Now we consider solutions in the bounded interval $\Omega = (0,\pi)$ with Dirichlet boundary conditions $\varphi(t,0) = \varphi(t,\pi) = 0$ corresponding to fixed homogeneous displacements on $\partial\Omega$. The solution can be expanded into eigenmodes of the operator $A\varphi = -\partial_x^2\varphi$ in the domain $\mathcal{D}(A) = \mathrm{H}_0^1(\Omega)$, so that we obtain

$$\varphi(t,x) = \sum_{k=1}^{\infty} \Big(\alpha_k \cos(ckt) + \beta_k \sin(ckt)\Big) \sin(kx) \,,$$

where the coefficients are determined by the initial displacement $\varphi(0,\cdot)$ and velocity $\partial_t\varphi(0,\cdot)$.

For the special example $\varphi(0,x) = 1$ and $\partial_t\varphi(0,x) = 0$ for $x \in (0,\pi)$ and $c = 1$, we obtain the explicit Fourier representation

$$\varphi(t,x) = \frac{4}{\pi}\sum_{k=0}^{\infty} \frac{1}{2k+1} \cos\big((2k+1)t\big) \sin\big((2k+1)x\big) = \frac{1}{2}\Big(\varphi_0(x+t) + \varphi_0(x-t)\Big)$$

with the periodic function $\varphi_0(x) = \begin{cases} 1 & x \in (0,\pi) + 2\pi\mathbb{Z} \,, \\ 0 & x \in \pi\mathbb{Z} \,, \\ -1 & x \in (-\pi,0) + 2\pi\mathbb{Z} \,. \end{cases}$

We observe that the solution is discontinuous along linear characteristics $x \pm ct = \mathrm{const}$.

In the space-time cylinder $Q = (0,T) \times \Omega$, the solution $\varphi(\cdot,\cdot)$ is a function in $\mathrm{BV}(Q)$.

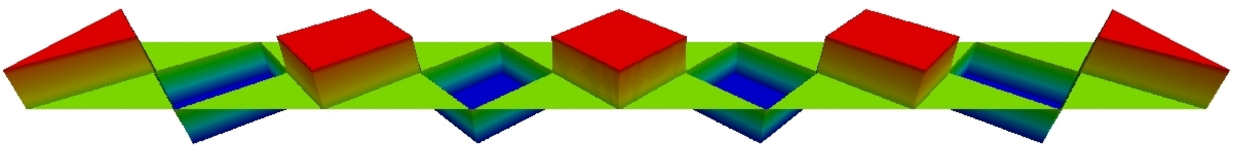


Figure 1. Solution φ in $(0,T) \times (0,\pi)$ with $T = 8$ for $\varphi(0,\cdot) = 1$, $\partial_t\varphi(0,\cdot) = 0$, and $\varphi(\cdot,0) = \varphi(\cdot,\pi) = 0$.

Waves in solids We consider the deformation vector $\boldsymbol{\varphi}(\cdot,\cdot)$ in an elastic solid $\Omega \subset \mathbb{R}^3$. The velocity is denoted by $\mathbf{v} = \partial_t \boldsymbol{\varphi}$. The elastic constitutive setting is determined by the stress response $\boldsymbol{\sigma} = \hat{\Sigma}(\mathbf{F})$ determining the stress tensor $\boldsymbol{\sigma}$ by a stress response function $\hat{\Sigma}(\cdot)$ from the deformation gradient $\mathbf{F} = \mathrm{D}\boldsymbol{\varphi}$. The stress rate is given by

$$\partial_t \boldsymbol{\sigma} = \mathrm{D}\hat{\Sigma}(\mathrm{D}\boldsymbol{\varphi})[\mathrm{D}\mathbf{v}]\,.$$

Assuming small strains and $\boldsymbol{\varphi} \approx \mathrm{id}$, this is approximated by its linearization

$$\partial_t \boldsymbol{\sigma} = \mathbf{C}[\mathrm{D}\mathbf{v}]$$

with the elasticity tensor $\mathbf{C} = \mathrm{D}\hat{\Sigma}(\mathbf{I})$. The balance of torsional moments yields that stress is symmetric and that the strain rate only depends on the symmetric strain rate $\boldsymbol{\varepsilon}(\mathbf{v}) = \mathrm{sym}(\mathrm{D}\mathbf{v})$. In isotropic media the elasticity tensor is characterized by the Lamé parameters $\lambda \geq 0$, $\mu > 0$, and introducing the compression modulus $\kappa = \frac{2\mu+3\lambda}{3}$ and the shear term $\mathrm{dev}(\boldsymbol{\sigma}) = \boldsymbol{\sigma} - \frac{1}{3}\,\mathrm{trace}(\boldsymbol{\sigma})\mathbf{I}$ we have

$$\mathbf{C}\boldsymbol{\varepsilon} = 2\mu\boldsymbol{\varepsilon} + \lambda\,\mathrm{trace}(\boldsymbol{\varepsilon})\mathbf{I} = 2\mu\,\mathrm{dev}(\boldsymbol{\varepsilon}) + \kappa\,\mathrm{trace}(\boldsymbol{\varepsilon})\mathbf{I}\,, \quad \mathbf{C}^{-1}\boldsymbol{\sigma} = \frac{1}{2\mu}\,\mathrm{dev}(\boldsymbol{\sigma}) + \frac{3}{\kappa}\,\mathrm{trace}(\boldsymbol{\sigma})\mathbf{I}.$$

Newton's law for the balance of momentum yields

$$\rho\partial_t \mathbf{v} = \mathrm{div}\,\boldsymbol{\sigma} + \mathbf{b}\,,$$

with the mass density ρ, acceleration $\partial_t \mathbf{v}$, and the vector of body forces $\mathbf{b}$.

Waves in compressible fluids In fluids we assume that the shear forces can be neglected, i.e., we consider the limit $\mu \longrightarrow 0$. Then, the stress $\boldsymbol{\sigma} = p\mathbf{I}$ is isotropic with hydrostatic pressure $p = \frac{1}{3}\,\mathrm{trace}\,\boldsymbol{\sigma}$, and compression waves are described by the system

$$\partial_t p = \kappa\,\mathrm{div}\,\mathbf{v}\,, \qquad \rho\partial_t \mathbf{v} = \nabla p + \mathbf{b}\,.$$

In particular this applies to acoustic waves in air or in a gas at fixed temperature.

Electro-magnetic waves By the laws of Ampere, Faraday, and Gauß we obtain the Maxwell system

$$\begin{array}{rclcrcl} \partial_t \mathbf{D} - \mathrm{curl}\,\mathbf{H} & = & -\mathbf{J}\,, & \qquad & \partial_t \mathbf{B} + \mathrm{curl}\,\mathbf{E} & = & \mathbf{0}\,, \\ \mathrm{div}\,\mathbf{D} & = & \rho\,, & & \mathrm{div}\,\mathbf{B} & = & 0 \end{array}$$

for the electric field $\mathbf{E}$, the magnetic field $\mathbf{H}$, the electric displacement $\mathbf{D}$, the magnetic field induction $\mathbf{B}$, the electric current density $\mathbf{J}$, and the electric charge density ρ. In vacuum with no electric charges, we have $\mathbf{J} = \mathbf{0}$, $\rho = 0$, and the material laws $\mathbf{D} = \varepsilon\mathbf{E}$ and $\mathbf{B} = \mu\mathbf{H}$ with permeability $\mu > 0$ and permittivity $\varepsilon > 0$.

First-order differential systems For all these cases we obtain a system of J equations in $\mathbb{R}^D$

$$M\partial_t \mathbf{u} + A\mathbf{u} = \mathbf{f}$$

with a first order differential operator A and a weighting operator M.

For *elastic waves*, we have $\mathbf{u} = (\boldsymbol{\sigma}, \mathbf{v})$, $A(\boldsymbol{\sigma}, \mathbf{v}) = -(\boldsymbol{\varepsilon}(\mathbf{v}), \mathrm{div}\,\boldsymbol{\sigma})$, and $M(\boldsymbol{\sigma}, \mathbf{v}) = (\mathbf{C}^{-1}\boldsymbol{\sigma}, \rho\mathbf{v})$, for *acoustic waves*, we have $\mathbf{u} = (p, \mathbf{v})$, $A(p, \mathbf{v}) = -(\mathrm{div}\,\mathbf{v}, \nabla p)$, and $M(p, \mathbf{v}) = (\kappa^{-1}p, \rho\mathbf{v})$, and for *electro-magnetic waves* $\mathbf{u} = (\mathbf{H}, \mathbf{E})$, $A(\mathbf{H}, \mathbf{E}) = (\mathrm{curl}\,\mathbf{E}, -\mathrm{curl}\,\mathbf{H})$, and $M(\mathbf{H}, \mathbf{E}) = (\mu\mathbf{H}, \varepsilon\mathbf{E})$.

2 A space-time setting for linear hyperbolic operators

Let $\Omega \subseteq \mathbb{R}^D$ be a bounded Lipschitz domain, and let $H \subseteq \mathrm{L}_2(\Omega;\mathbb{R}^J)$ be a Hilbert space with weighted inner product $(\mathbf{v},\mathbf{w})_H = (M\mathbf{v},\mathbf{w})_{0,\Omega}$, where $M \in \mathrm{L}_\infty(\Omega,\mathbb{R}^{J\times J}_{\mathsf{sym}})$ is uniformly positive. We consider a linear operator in space $A \in \mathcal{L}(\mathcal{D}(A),H)$ with domain $\mathcal{D}(A) \subset H$.

Homogeneous boundary conditions are defined by the domain. Here, we choose the domain $\mathcal{D}(A) = \mathrm{H}(\mathrm{div},\Omega;\mathbb{R}^{D\times D}_{\mathsf{sym}}) \times \mathrm{H}^1_0(\Omega;\mathbb{R}^D)$ for elastic waves, $\mathcal{D}(A) = \mathrm{H}^1(\Omega) \times \mathrm{H}_0(\mathrm{div},\Omega)$ for acoustic waves, and for electro-magnetic waves we choose $\mathcal{D}(A) = \mathrm{H}_0(\mathrm{curl},\Omega) \times \mathrm{H}(\mathrm{curl},\Omega)$. In all cases the operator is skew-adjoint, i.e.,

$$(A\mathbf{v},\mathbf{w})_{0,\Omega} = -(\mathbf{v},A\mathbf{w})_{0,\Omega}\,, \qquad \mathbf{v},\mathbf{w} \in \mathcal{D}(A)\,.$$

The variational setting In the abstract setting, we consider the operator $L = M\partial_t + A$ on the space-time cylinder $Q = \Omega \times (0,T)$. Again, we observe

$$(L\mathbf{v},\mathbf{w})_{0,Q} = -(\mathbf{v},L\mathbf{w})_{0,Q}\,, \qquad \mathbf{v},\mathbf{w} \in \mathrm{C}^1_0(Q;\mathbb{R}^J)\,.$$

Depending on L we define the space

$$\begin{aligned} \mathrm{H}(L,Q) \quad = \quad & \big\{\mathbf{v} \in \mathrm{L}_2(Q;\mathbb{R}^J)\colon \mathbf{g} \in \mathrm{L}_2(Q;\mathbb{R}^J) \text{ exists with} \\ & \qquad\qquad (\mathbf{g},\mathbf{w})_{0,Q} = -(\mathbf{v},L\mathbf{w})_{0,Q} \text{ for all } \mathbf{w} \in \mathrm{C}^1_0(Q;\mathbb{R}^J)\big\}\,. \end{aligned}$$

Then, L can be extended to this space, and $\mathrm{H}(L,\Omega)$ is a Hilbert space with respect to the weighted graph norm $\|\mathbf{v}\|_{L,Q} = \sqrt{\|\mathbf{v}\|^2_{0,Q} + \|L\mathbf{v}\|^2_{0,Q}}$.

Let $V \subset \mathrm{H}(L,Q)$ be the closure of $\big\{\mathbf{v} \in \mathrm{C}^1([0,T];\mathcal{D}(A))\colon \mathbf{v}(0) = \mathbf{0}\big\}$. Then we define $W = \overline{L(V)} \subseteq \mathrm{L}_2(Q;\mathbb{R}^J)$ with the weighted norm $\|\mathbf{w}\|^2_W = (M\mathbf{w},\mathbf{w})_{0,Q}$. On V, we use the weighted graph norm $\|\mathbf{v}\|^2_V = \|\mathbf{v}\|^2_W + \|M^{-1}L\mathbf{v}\|^2_W$.

Now we study the operator L in $\mathcal{L}(V,W)$ and the evolution equation $L\mathbf{u} = \mathbf{f}$ with homogeneous initial and boundary conditions. This extends to initial values $\mathbf{u}_0 \neq \mathbf{0}$ by replacing $\mathbf{f}(t)$ with $\mathbf{f}(t) - A\mathbf{u}_0$. Also inhomogeneous boundary conditions can be analyzed by modifying the right-hand side when the existence of a sufficiently smooth extension of the boundary data can be assumed.

We define the bilinear form $b\colon V \times W \longrightarrow \mathbb{R}$ with $b(\mathbf{v},\mathbf{w}) = (L\mathbf{v},\mathbf{w})_{0,Q}$, and we establish the standard Babuška setting (see, e.g., [Bra07, Thm. III.3.6]).

Lemma 2.1 *Assume that $(A\mathbf{z},\mathbf{z})_{0,\Omega} \geq 0$ for $\mathbf{z} \in D(A)$. Then, the bilinear form $b(\cdot,\cdot)$ is continuous and inf-sup stable in $V \times W$ with $\beta = (4T^2+1)^{-1/2}$, i.e.,*

$$\sup_{\mathbf{w}\in W\setminus\{\mathbf{0}\}} \frac{b(\mathbf{v},\mathbf{w})}{\|\mathbf{w}\|_W} \geq \beta\,\|\mathbf{v}\|_V\,, \qquad \mathbf{v} \in V\,.$$

Proof. The continuity follows from the upper bound $|b(\mathbf{v},\mathbf{w})| \leq \|\mathbf{v}\|_V\|\mathbf{w}\|_W$. To prove the inf-sup condition we first note that for all $\mathbf{v} \in \mathrm{C}^1([0,T];\mathcal{D}(A))$ with $\mathbf{v}(0) = \mathbf{0}$ we have

$$\begin{aligned} \|\mathbf{v}\|^2_W &= \int_0^T \big(M\mathbf{v}(t),\mathbf{v}(t)\big)_{0,\Omega}\,\mathrm{d}t = \int_0^T \Big(\big(M\mathbf{v}(t),\mathbf{v}(t)\big)_{0,\Omega} - \big(M\mathbf{v}(0),\mathbf{v}(0)\big)_{0,\Omega}\Big)\,\mathrm{d}t \\ &= \int_0^T\int_0^t \partial_t\big(M\mathbf{v}(s),\mathbf{v}(s)\big)_{0,\Omega}\,\mathrm{d}s\,\mathrm{d}t = 2\int_0^T\int_0^t \big(M\partial_t\mathbf{v}(s),\mathbf{v}(s)\big)_{0,\Omega}\,\mathrm{d}s\,\mathrm{d}t \\ &\leq 2\int_0^T\int_0^t \big(M\partial_t\mathbf{v}(s) + A\mathbf{v}(s),\mathbf{v}(s)\big)_{0,\Omega}\,\mathrm{d}s\,\mathrm{d}t \\ &\leq 2\int_0^T\int_0^t \big(M^{-1}L\mathbf{v}(s),L\mathbf{v}(s)\big)^{1/2}_{0,\Omega}\big(M\mathbf{v}(s),\mathbf{v}(s)\big)^{1/2}_{0,\Omega}\,\mathrm{d}s\,\mathrm{d}t \leq 2T\,\|M^{-1}L\mathbf{v}\|_W\|\mathbf{v}\|_W\,. \end{aligned}$$

This yields $\|\mathbf{v}\|_W \le 2T\,\|M^{-1}L\mathbf{v}\|_W$ for $\mathbf{v}\in V$. Let $\mathbf{v}\in V\setminus\{\mathbf{0}\}$ and take $\mathbf{w}=M^{-1}L\mathbf{v}\in W\setminus\{\mathbf{0}\}$, then

$$\sup_{\mathbf{w}\in W\setminus\{\mathbf{0}\}}\frac{b(\mathbf{v},\mathbf{w})}{\|\mathbf{w}\|_W}\ge\frac{b(\mathbf{v},M^{-1}L\mathbf{v})}{\|M^{-1}L\mathbf{v}\|_W}=\frac{(L\mathbf{v},M^{-1}L\mathbf{v})_{0,\Omega}}{\|M^{-1}L\mathbf{v}\|_W}=\|M^{-1}L\mathbf{v}\|_W\ge\frac{1}{\sqrt{4T^2+1}}\,\|\mathbf{v}\|_V\,,$$

where the final inequality follows from $\|\mathbf{v}\|_V^2=\|\mathbf{v}\|_W^2+\|M^{-1}L\mathbf{v}\|_W^2\le(4T^2+1)\|M^{-1}L\mathbf{v}\|_W^2$. $\square$

The inf-sup stability ensures that the operator $L\in\mathcal{L}(V,W)$ is injective and that the range is closed. Thus, the operator is surjective by construction and the inverse L^{-1} is bounded in $\mathcal{L}(W,V)$. This yields directly the following result [Bra07, Thm. III.3.6].

Theorem 2.2 *For given $\mathbf{f}\in\mathrm{L}_2(Q;\mathbb{R}^J)$ there exists a unique solution $\mathbf{u}\in V$ of*

$$(L\mathbf{u},\mathbf{w})_{0,Q}=(\mathbf{f},\mathbf{w})_{0,Q}\,,\qquad \mathbf{w}\in W \tag{1}$$

satisfying the a priori bound $\|\mathbf{u}\|_V\le\sqrt{4T^2+1}\,\|M^{-1/2}\mathbf{f}\|_{0,Q}$.

Remark 2.3 *The approach presented here to show that $L\in\mathcal{L}(V,W)$ is an invertible operator in suitable Hilbert space V and W only requires to show that L is injective. Since L mixes the derivatives in space and time, more regularity is difficult to show in this framework. Therefore, one can check the assumptions of the Lumer-Phillips theorem [RR04, Chap. 12.2.2] for the operator A in $\mathcal{D}(A)$, so that semigroup theory with more regularity can applied, see, e.g., [EN06].*

3 Discontinuous Galerkin methods for linear systems of conservation laws

All wave equations discussed above can be considered as a system of linear conservation laws

$$M\partial_t \mathbf{u}(t) + \operatorname{div} \mathbf{F}(\mathbf{u}(t)) = \mathbf{f}(t) \quad \text{for} \quad t \in [0,T]\,, \qquad \mathbf{u}(0) = \mathbf{u}_0 \tag{2}$$

with a linear *flux function* $\mathbf{F}(\mathbf{v}) = [B_1\mathbf{v}, \dots, B_D\mathbf{v}]$ and symmetric matrices $B_d \in \mathbb{R}^{J\times J}_{\mathsf{sym}}$ such that

$$A\mathbf{v} = \operatorname{div} \mathbf{F}(\mathbf{v}) = \sum_{d=1}^{D} B_d \partial_d \mathbf{v}\,.$$

Traveling waves In the case of constant coefficients in $\Omega = \mathbb{R}^D$, special solutions can be constructed as follows. For a given unit vector $\mathbf{n} = (n_1, \dots, n_D)^T \in \mathbb{R}^D$, we have $\mathbf{n}\cdot\mathbf{F}(\mathbf{u}) = B\mathbf{u}$ with the symmetric matrix $B = \sum n_d B_d$. Then, for all eigenpairs $(\lambda, \mathbf{w}) \in \mathbb{R}\times\mathbb{R}^J$ with $B\mathbf{w} = \lambda M\mathbf{w}$ and amplitude function $a(\cdot)$, the traveling wave propagating with velocity $c = |\lambda|$

$$\mathbf{u}(t,\mathbf{x}) = a(\mathbf{n}\cdot\mathbf{x} - \lambda t)\,\mathbf{w}$$

is a solution of (2) with initial value $\mathbf{u}_0(\mathbf{x}) = a(\mathbf{n}\cdot\mathbf{x})\,\mathbf{w}$ and right-hand side $\mathbf{f} = \mathbf{0}$.

Discontinuous weak solutions A function $\mathbf{u} \in \mathrm{L}_1((0,T)\times\mathbb{R}^D;\mathbb{R}^J)$ is a weak solution of (2) if

$$0 = \int_{\mathbb{R}^D} M(\mathbf{x})\mathbf{u}_0(\mathbf{x})\cdot\boldsymbol{\phi}(0,\mathbf{x})\,\mathrm{d}\mathbf{x} + \int_{(0,T)\times\mathbb{R}^D} \mathbf{u}(t,\mathbf{x})\cdot\Big(M(\mathbf{x})\partial_t\boldsymbol{\phi}(t,\mathbf{x}) + \operatorname{div}\mathbf{F}\big(\boldsymbol{\phi}(t,\mathbf{x})\big) - \mathbf{f}(t,\mathbf{x})\cdot\boldsymbol{\phi}(t,\mathbf{x})\Big)\,\mathrm{d}t\,\mathrm{d}\mathbf{x}$$

for all test functions with compact support $\boldsymbol{\phi} \in \mathrm{C}^1_0((-1,T)\times\Omega;\mathbb{R}^J)$.

This applies to traveling waves with discontinuous amplitude: the piecewise constant function

$$\mathbf{u}(t,\mathbf{x}) = \begin{cases} a_{\mathsf{L}}\mathbf{w} & \text{in } Q_{\mathsf{L}} = \big\{(t,\mathbf{x}) \in [0,T]\times\mathbb{R}^D \colon \mathbf{n}\cdot\mathbf{x} - \lambda t < 0\big\} \\ a_{\mathsf{R}}\mathbf{w} & \text{in } Q_{\mathsf{R}} = \big\{(t,\mathbf{x}) \in [0,T]\times\mathbb{R}^D \colon \mathbf{n}\cdot\mathbf{x} - \lambda t > 0\big\} \end{cases} \tag{3}$$

with $a_{\mathsf{L}}, a_{\mathsf{R}} \in \mathbb{R}$ is a weak solution: Using $(-\lambda M + B)\mathbf{w} = \mathbf{0}$ and the Gauß theorem in $Q_{\mathsf{L}} \subset \mathbb{R}\times\mathbb{R}^D$ with unit normal vector $\frac{1}{\sqrt{1+\lambda^2}}\begin{pmatrix} -\lambda \\ \mathbf{n} \end{pmatrix}$ on ∂Q_{L}, we observe

$$\begin{aligned}
0 &= a_{\mathsf{L}} \int_{\mathbf{n}\cdot\mathbf{x}-\lambda t=0} \big(-\lambda M + B\big)\mathbf{w}\cdot\boldsymbol{\phi}(t,\mathbf{x})\,\mathrm{d}\mathbf{a} \\
&= \int_{\partial Q_{\mathsf{L}}} \begin{pmatrix} -\lambda \\ \mathbf{n} \end{pmatrix} \cdot \begin{pmatrix} M\mathbf{u}(t,\mathbf{x})\cdot\boldsymbol{\phi}(t,\mathbf{x}) \\ \mathbf{F}\big(\mathbf{u}(t,x)\big)\cdot\boldsymbol{\phi}(t,\mathbf{x}) \end{pmatrix}\,\mathrm{d}\mathbf{a} \\
&= \sqrt{1+\lambda^2}\int_{Q_{\mathsf{L}}} \begin{pmatrix} \partial_t \\ \nabla \end{pmatrix} \cdot \begin{pmatrix} \mathbf{u}(t,\mathbf{x})\cdot M\boldsymbol{\phi}(t,\mathbf{x}) \\ \mathbf{u}(t,x)\cdot\mathbf{F}\big(\boldsymbol{\phi}(t,\mathbf{x})\big) \end{pmatrix}\,\mathrm{d}t\,\mathrm{d}\mathbf{x} \\
&= \sqrt{1+\lambda^2}\int_{Q_{\mathsf{L}}} \Big(\mathbf{u}(t,\mathbf{x})\cdot M\partial_t\boldsymbol{\phi}(t,\mathbf{x}) + \mathbf{u}(t,x)\cdot\operatorname{div}\mathbf{F}\big(\boldsymbol{\phi}(t,\mathbf{x})\big)\Big)\,\mathrm{d}t\,\mathrm{d}\mathbf{x}
\end{aligned}$$

for all test functions $\boldsymbol{\phi} \in \mathrm{C}^1_0\big((0,T)\times\mathbb{R}^D;\mathbb{R}^J\big)$. Repeating this argument with ∂Q_{R} and testing in $\mathrm{C}^1_0\big((-1,T)\times\mathbb{R}^D;\mathbb{R}^J\big)$ shows that (3) is a weak solution with discontinuity along the hyperplane

$$\partial Q_{\mathsf{R}} = \partial Q_{\mathsf{L}} = \{(t,\mathbf{x})\colon \mathbf{n}\cdot\mathbf{x} - \lambda t = 0\}$$

in the time-space cylinder and with discontinuous initial values $\mathbf{u}_0(\mathbf{x}) = \mathbf{u}(0,\mathbf{x})$.

The Riemann problem for linear conservation laws We now construct a weak solution of the Riemann problem, i.e., a piecewise constant weak solution with right-hand side $\mathbf{f} = \mathbf{0}$ and the discontinuous initial function

$$\mathbf{u}_0(\mathbf{x}) \quad = \quad \begin{cases} \mathbf{u}_\mathsf{L} & \text{in } \Omega_\mathsf{L} = \big\{\mathbf{x} \in \mathbb{R}^D \colon \mathbf{n}\cdot\mathbf{x} < 0\big\} \\ \mathbf{u}_\mathsf{R} & \text{in } \Omega_\mathsf{R} = \big\{\mathbf{x} \in \mathbb{R}^D \colon \mathbf{n}\cdot\mathbf{x} > 0\big\} \end{cases} \tag{4}$$

with $\mathbf{u}_\mathsf{L}, \mathbf{u}_\mathsf{R} \in \mathbb{R}^J$. Let $(\lambda_j, \mathbf{w}_j)$ be M-orthogonal eigenpairs, i.e.,

$$B\mathbf{w}_j = \lambda_j M\mathbf{w}_j \text{ with } \mathbf{w}_k \cdot M\mathbf{w}_j = 0 \text{ for } j \neq k\,.$$

By the superposition of traveling waves, we obtain a weak solution of the Riemann problem

$$\mathbf{u}(t,\mathbf{x}) \quad = \quad \sum_j a_j(\mathbf{x}\cdot\mathbf{n} - \lambda_j t)\mathbf{w}_j\,, \quad a_j(s) = \begin{cases} \mathbf{w}_j \cdot M\mathbf{u}_\mathsf{L}/\mathbf{w}_j \cdot M\mathbf{w}_j & s < 0\,, \\ \mathbf{w}_j \cdot M\mathbf{u}_\mathsf{R}/\mathbf{w}_j \cdot M\mathbf{w}_j & s > 0\,. \end{cases} \tag{5}$$

The solution of the Riemann problem at $(t,\mathbf{0})$ for $t > 0$ defines the upwind flux on $\partial\Omega_\mathsf{L} \cap \partial\Omega_\mathsf{R}$ by

$$\mathbf{n}\cdot\mathbf{F}^{\mathrm{up}}(\mathbf{u}_0) = \sum_{\lambda_j>0} \frac{\mathbf{w}_j \cdot M\mathbf{u}_\mathsf{L}}{\mathbf{w}_j \cdot M\mathbf{w}_j}\, B\mathbf{w}_j + \sum_{\lambda_j<0} \frac{\mathbf{w}_j \cdot M\mathbf{u}_\mathsf{R}}{\mathbf{w}_j \cdot M\mathbf{w}_j}\, B\mathbf{w}_j = B\mathbf{u}_\mathsf{L} + \sum_{\lambda_j<0} \frac{\mathbf{w}_j \cdot B[\mathbf{u}]}{\mathbf{w}_j \cdot B\mathbf{w}_j}\, B\mathbf{w}_j \tag{6}$$

depending on the jump term $[\mathbf{u}] = \mathbf{u}_\mathsf{R} - \mathbf{u}_\mathsf{L}$. By construction, the upwind flux is consistent, i.e., for $B\mathbf{u}_\mathsf{L} = B\mathbf{u}_\mathsf{R}$ we obtain $\mathbf{n}\cdot\mathbf{F}^{\mathrm{up}}(\mathbf{u}_0) = B\mathbf{u}_\mathsf{L} = B\mathbf{u}_\mathsf{R}$.

Application to wave equations For *elastic waves* with $\operatorname{div}\mathbf{F}(\boldsymbol{\sigma},\mathbf{v}) = -\begin{pmatrix}\boldsymbol{\varepsilon}(\mathbf{v}) \\ \operatorname{div}\boldsymbol{\sigma}\end{pmatrix}$ we have the normal flux $\mathbf{n}\cdot\mathbf{F}(\boldsymbol{\sigma},\mathbf{v}) = -\begin{pmatrix}\frac12(\mathbf{n}\otimes\mathbf{v} + \mathbf{v}\otimes\mathbf{n}) \\ \boldsymbol{\sigma}\mathbf{n}\end{pmatrix}$. By $c_\mathsf{P} = \sqrt{\frac{2\mu+\lambda}{\rho}}$ we denote the velocity of pressure waves, and by $c_\mathsf{S} = \sqrt{\frac{\mu}{\rho}}$ the velocity of shear waves. The eigenvectors are of the form $\begin{pmatrix}2\mu\mathbf{n}\otimes\mathbf{n} + \lambda\mathbf{I} \\ \pm c_\mathsf{P}\mathbf{n}\end{pmatrix}$ and $\begin{pmatrix}\mu(\boldsymbol{\tau}\otimes\mathbf{n} + \mathbf{n}\otimes\boldsymbol{\tau}) \\ \pm c_\mathsf{S}\boldsymbol{\tau}\end{pmatrix}$, where $\boldsymbol{\tau}$ is a unit tangent vector, i.e., $\boldsymbol{\tau}\cdot\mathbf{n} = 0$. The resulting upwind flux in 2D is given by

$$\begin{aligned}\mathbf{n}\cdot\mathbf{F}^{\mathrm{up}}(\mathbf{u}_0) \quad = \quad & -\begin{pmatrix}\frac12(\mathbf{n}\otimes\mathbf{v}_\mathsf{L} + \mathbf{v}_\mathsf{L}\otimes\mathbf{n}) \\ \boldsymbol{\sigma}_\mathsf{L}\mathbf{n}\end{pmatrix} - \frac{\mathbf{n}\cdot[\boldsymbol{\sigma}]\mathbf{n} + \rho c_\mathsf{P}[\mathbf{v}]\cdot\mathbf{n}}{2\rho c_\mathsf{P}}\begin{pmatrix}\mathbf{n}\otimes\mathbf{n} \\ \rho c_\mathsf{P}\mathbf{n}\end{pmatrix} \\ & - \frac{\boldsymbol{\tau}\cdot[\boldsymbol{\sigma}]\mathbf{n} + \rho c_\mathsf{S}[\mathbf{v}]\cdot\boldsymbol{\tau}}{2\rho c_\mathsf{S}}\begin{pmatrix}\frac12(\boldsymbol{\tau}\otimes\mathbf{n} + \mathbf{n}\otimes\boldsymbol{\tau}) \\ \rho c_\mathsf{S}\boldsymbol{\tau}\end{pmatrix}\,.\end{aligned}$$

For *acoustic waves* with $\operatorname{div}\mathbf{F}(p,\mathbf{v}) = -\begin{pmatrix}\operatorname{div}\mathbf{v} \\ \nabla p\end{pmatrix}$ we obtain $\mathbf{n}\cdot\mathbf{F}(p,\mathbf{v}) = -\begin{pmatrix}\mathbf{v}\cdot\mathbf{n} \\ p\mathbf{n}\end{pmatrix}$, the velocity of sound $c = \sqrt{\frac{\kappa}{\rho}}$, and the eigenvectors $\mathbf{w}_\pm = \begin{pmatrix}\kappa \\ \mp c\mathbf{n}\end{pmatrix}$. This yields the upwind flux

$$\mathbf{n}\cdot\mathbf{F}^{\mathrm{up}}(\mathbf{u}_0) \quad = \quad -\begin{pmatrix}\mathbf{v}_\mathsf{L}\cdot\mathbf{n} \\ p_\mathsf{L}\mathbf{n}\end{pmatrix} - \frac{[p] + \rho c\mathbf{n}\cdot[\mathbf{v}]}{2\rho c}\begin{pmatrix}1 \\ \rho c\mathbf{n}\end{pmatrix}\,.$$

For *electro-magnetic waves* with $\operatorname{div}\mathbf{F}(\mathbf{H},\mathbf{E}) = \begin{pmatrix}\operatorname{curl}\mathbf{E} \\ -\operatorname{curl}\mathbf{H}\end{pmatrix}$ and $\mathbf{n}\cdot\mathbf{F}(\mathbf{H},\mathbf{E}) = \begin{pmatrix}\mathbf{n}\times\mathbf{E} \\ -\mathbf{n}\times\mathbf{H}\end{pmatrix}$ we have the speed of light $c = \frac{1}{\sqrt{\varepsilon\mu}}$, eigenvectors $\begin{pmatrix}\sqrt{\varepsilon}\mathbf{n}\times\boldsymbol{\tau} \\ \pm\sqrt{\mu}\boldsymbol{\tau}\end{pmatrix}$ and $\begin{pmatrix}\mp\sqrt{\varepsilon}\,\boldsymbol{\tau} \\ \sqrt{\mu}\,\mathbf{n}\times\boldsymbol{\tau}\end{pmatrix}$, and upwind flux

$$\begin{aligned}\mathbf{n}\cdot\mathbf{F}^{\mathrm{up}}(\mathbf{u}_0) \quad = \quad & \begin{pmatrix}\mathbf{n}\times\mathbf{E}_\mathsf{L} \\ -\mathbf{n}\times\mathbf{H}_\mathsf{L}\end{pmatrix} + \frac{\sqrt{\mu}[\mathbf{H}]\cdot(\mathbf{n}\times\boldsymbol{\tau}) - \sqrt{\varepsilon}[\mathbf{E}]\cdot\boldsymbol{\tau}}{2\sqrt{\mu\varepsilon}}\begin{pmatrix}-\sqrt{\mu}\mathbf{n}\times\boldsymbol{\tau} \\ \sqrt{\varepsilon}\boldsymbol{\tau}\end{pmatrix} \\ & + \frac{\sqrt{\mu}[\mathbf{H}]\cdot\boldsymbol{\tau} + \sqrt{\varepsilon}[\mathbf{E}]\cdot(\mathbf{n}\times\boldsymbol{\tau})}{2\sqrt{\mu\varepsilon}}\begin{pmatrix}\sqrt{\mu}\boldsymbol{\tau} \\ \sqrt{\varepsilon}\mathbf{n}\times\boldsymbol{\tau}\end{pmatrix}\,.\end{aligned}$$

The discontinuous Galerkin discretization in space We assume that Ω is a bounded polyhedral Lipschitz domain decomposed into a finite number of open elements $K \subset \Omega$ such that $\overline{\Omega} = \bigcup_{K\in\mathcal{K}} \overline{K}$, where $\mathcal{K}$ is the set of elements in space. Let $\mathcal{F}_K$ be the set of faces of K, and for inner faces $f \in \mathcal{F}_K$ let K_f be the neighboring cell such that $f = \partial K \cap \partial K_f$, and let $\mathbf{n}_K$ be the outer unit normal vector on ∂K. The outer unit normal vector field on $\partial\Omega$ is denoted by $\mathbf{n}$.

We select polynomial degrees p_K, and we define the local spaces $H_{h,K} = \mathbb{P}_{p_K}(K;\mathbb{R}^J)$ and the global discontinuous Galerkin space

$$H_h = \{\mathbf{v}_h \in \mathrm{L}_2(\Omega)^J \colon \mathbf{v}_h|_K \in H_{h,K} \text{ for all } K \in \mathcal{K}\}.$$

For $\mathbf{v}_h \in H_h$ we define $\mathbf{v}_{h,K} = \mathbf{v}_h|_K \in H_{h,K}$ for the restriction to K. In the semi-discrete problem

$$M_h\partial_t\mathbf{u}_h(t) + A_h\mathbf{u}_h(t) = \mathbf{f}(t)\,, \qquad t \in (0,T)\,, \tag{7}$$

the discrete mass operator $M_h \in \mathcal{L}(H_h, H_h)$ is the Galerkin approximation of M defined by

$$(M_h\mathbf{v}_h, \mathbf{w}_h)_{0,\Omega} = (M\mathbf{v}_h, \mathbf{w}_h)_{0,\Omega} \qquad \mathbf{v}_h, \mathbf{w}_h \in H_h\,. \tag{8}$$

The discrete mass operator M_h is represented by a block diagonal positive definite matrix.

The discrete operator $A_h \in \mathcal{L}(H_h, H_h)$ is constructed as follows: Integration by parts yields for smooth ansatz functions $\mathbf{v}$ and smooth test functions $\boldsymbol{\phi}_K$

$$(A\mathbf{v}, \boldsymbol{\phi}_K)_{0,K} = (\operatorname{div}\mathbf{F}(\mathbf{v}), \boldsymbol{\phi}_K)_{0,K} = -(\mathbf{F}(\mathbf{v}), \nabla\boldsymbol{\phi}_K)_{0,K} + \sum_{f\in\mathcal{F}_K} (\mathbf{n}_K\cdot\mathbf{F}(\mathbf{v}), \boldsymbol{\phi}_K)_{0,f}\,.$$

We then define for $\mathbf{v}_h \in H_h$ and $\boldsymbol{\phi}_{h,K} \in H_{h,K}$ by

$$(A_h\mathbf{v}_h, \boldsymbol{\phi}_{h,K})_{0,K} = -(\mathbf{F}(\mathbf{v}_{h,K}), \nabla\boldsymbol{\phi}_{h,K})_{0,K} + \sum_{f\in\mathcal{F}_K} \big(\mathbf{n}_K\cdot\mathbf{F}_K^{\mathrm{up}}(\mathbf{v}_h), \boldsymbol{\phi}_{h,K}\big)_{0,f}\,,$$

where $\mathbf{n}_K\cdot\mathbf{F}_K^{\mathrm{up}}(\mathbf{v}_h)$ is the upwind flux obtained from local solutions of Riemann problems. Again using integration by parts, we obtain

$$(A_h\mathbf{v}_h, \boldsymbol{\phi}_{h,K})_{0,K} = \big(\operatorname{div}\mathbf{F}(\mathbf{v}_{h,K}), \boldsymbol{\phi}_{h,K}\big)_{0,K} + \sum_{f\in\mathcal{F}_K} \big(\mathbf{n}_K\cdot(\mathbf{F}_K^{\mathrm{up}}(\mathbf{v}_h) - \mathbf{F}(\mathbf{v}_{h,K})), \boldsymbol{\phi}_{h,K}\big)_{0,f}\,. \tag{9}$$

On inner faces $f = \partial K \cap \partial K_f$ the difference $\mathbf{n}_K\cdot(\mathbf{F}_K^{\mathrm{up}}(\mathbf{v}_h) - \mathbf{F}(\mathbf{v}_{h,K}))$ only depends on the jump term $[\mathbf{v}_h]_{K,f} = \mathbf{v}_{h,K_f} - \mathbf{v}_{h,K}$, so that $\mathbf{n}_K\cdot(\mathbf{F}_K^{\mathrm{up}}(\mathbf{v}) - \mathbf{F}(\mathbf{v})) = 0$ on all faces $f \in \mathcal{F}_K$ for $\mathbf{v} \in D(A)$. On boundary faces, we define the jump term $[\mathbf{v}_h]_{K,f}$ depending on the boundary conditions. Together, we obtain consistency

$$(A\mathbf{v}, \boldsymbol{\phi}_h)_{0,\Omega} = (A_h\mathbf{v}, \boldsymbol{\phi}_h)_{0,\Omega}\,, \qquad \mathbf{v} \in D(A)\,,\ \boldsymbol{\phi}_h \in H_h\,, \tag{10}$$

and

$$\sum_{K\in\mathcal{K}} \big(\mathbf{n}_K\cdot\mathbf{F}_K^{\mathrm{up}}(\mathbf{v}_{h,K}), \mathbf{v}\big)_{0,\partial K} = 0\,, \qquad \mathbf{v}_h \in H_h\,,\ \mathbf{v} \in D(A)\cap\mathrm{H}^1(\Omega;\mathbb{R}^J)\,. \tag{11}$$

The upwind flux together with the appropriate choice of boundary flux guarantees that the discrete operator is non-negative and controls the nonconformity, i.e., a constant $C_{\mathsf{A}} > 0$ exists such that

$$(A_h\mathbf{v}_h, \mathbf{v}_h)_{0,\Omega} \ge C_{\mathsf{A}} \sum_{f\in\mathcal{F}_K} \big\|\big(\mathbf{n}_K\cdot(\mathbf{F}_K^{\mathrm{up}}(\mathbf{v}_h) - \mathbf{F}(\mathbf{v}_{h,K}))\big\|_{0,f}^2 \ge 0\,, \qquad \mathbf{v}_h \in H_h\,. \tag{12}$$

This is now shown for all our applications.

For *elastic waves* we obtain for $(\boldsymbol{\sigma}_h, \mathbf{v}_h) \in V_h$ and $(\boldsymbol{\varphi}_{K,h}, \boldsymbol{\psi}_{K,h}) \in V_{K,h}$

$$\begin{aligned}\big(A_h(\boldsymbol{\sigma}_h, \mathbf{v}_h), (\boldsymbol{\varphi}_{K,h}, \boldsymbol{\psi}_{K,h})\big)_{0,K} = &-\big(\boldsymbol{\varepsilon}(\mathbf{v}_{K,h}), \boldsymbol{\varphi}_{K,h}\big)_{0,K} - \big(\operatorname{div} \boldsymbol{\sigma}_{K,h}, \boldsymbol{\psi}_{K,h}\big)_{0,K}\\ &-\frac{1}{2\rho c_{\mathsf{S}}} \sum_{f \in \mathcal{F}_K} \big(\mathbf{n}_K \times ([\boldsymbol{\sigma}]_{K,f}\mathbf{n}_K + \rho c_{\mathsf{P}}[\mathbf{v}]_{K,f}), \mathbf{n}_K \times (\boldsymbol{\varphi}_{K,h}\mathbf{n}_K + \rho c_{\mathsf{P}}\boldsymbol{\psi}_{K,h})\big)_{0,f}\\ &-\frac{1}{2\rho c_{\mathsf{P}}} \sum_{f \in \mathcal{F}_K} \big(\mathbf{n}_K \cdot ([\boldsymbol{\sigma}]_{K,f}\mathbf{n}_K + \rho c_{\mathsf{P}}[\mathbf{v}]_{K,f}), \mathbf{n}_K \cdot (\boldsymbol{\varphi}_{K,h}\mathbf{n}_K + \rho c_{\mathsf{P}}\boldsymbol{\psi}_{K,h})\big)_{0,f} .\end{aligned}$$

On boundary faces $f = \partial K \cap \partial\Omega$, we set $[\mathbf{v}_h]_{K,f} = -2\mathbf{v}_{K,h}$ and $[\boldsymbol{\sigma}_h]_{K,f} = \mathbf{0}$ for Dirichlet boundary conditions. This yields

$$\begin{aligned}\big(A_h(\boldsymbol{\sigma}_h, \mathbf{v}_h), (\boldsymbol{\sigma}_{K,h}, \mathbf{v}_{K,h})\big)_{0,K} = &\sum_K \sum_{f \in \mathcal{F}_K} \Big(- \big(\mathbf{v}_{K,h}, \boldsymbol{\sigma}_{K,h}\mathbf{n}_K\big)_{0,f}\\ &-\frac{1}{2\rho c_{\mathsf{S}}} \sum_{f \in \mathcal{F}_K} \big(\mathbf{n}_K \times ([\boldsymbol{\sigma}]_{K,f}\mathbf{n}_K + \rho c_{\mathsf{P}}[\mathbf{v}]_{K,f}), \mathbf{n}_K \times (\boldsymbol{\sigma}_{K,h}\mathbf{n}_K + \rho c_{\mathsf{P}}\mathbf{v}_{K,h})\big)_{0,f}\\ &-\frac{1}{2\rho c_{\mathsf{P}}} \sum_{f \in \mathcal{F}_K} \big(\mathbf{n}_K \cdot ([\boldsymbol{\sigma}]_{K,f}\mathbf{n}_K + \rho c_{\mathsf{P}}[\mathbf{v}]_{K,f}), \mathbf{n}_K \cdot (\boldsymbol{\sigma}_{K,h}\mathbf{n}_K + \rho c_{\mathsf{P}}\mathbf{v}_{K,h})\big)_{0,f}\Big)\\ = &\frac{1}{2} \sum_K \sum_{f \in \mathcal{F}_K} \Big(\frac{1}{\rho c_{\mathsf{S}}} \big\|\mathbf{n}_K \times [\boldsymbol{\sigma}]_{K,f}\mathbf{n}_K\big\|_{0,f}^2 + \rho c_{\mathsf{S}} \big\|\mathbf{n}_K \times [\mathbf{v}_h]_{K,f}\big\|_{0,f}^2\\ &+\frac{1}{\rho c_{\mathsf{P}}} \big\|\mathbf{n}_K \cdot [\boldsymbol{\sigma}]_{K,f}\mathbf{n}_K\big\|_{0,f}^2 + \rho c_{\mathsf{P}} \big\|\mathbf{n}_K \cdot [\mathbf{v}_h]_{K,f}\big\|_{0,f}^2\Big) .\end{aligned}$$

For *acoustic waves* we obtain for $(p_h, \mathbf{v}_h) \in V_h$ and $(\varphi_{K,h}, \boldsymbol{\psi}_{K,h}) \in V_{K,h}$

$$\begin{aligned}\big(A_h(p_h, \mathbf{v}_h), (\varphi_{K,h}, \boldsymbol{\psi}_{K,h})\big)_{0,K} \quad = \quad &-\big(\operatorname{div} \mathbf{v}_{K,h}, \varphi_{K,h}\big)_{0,K} - \big(\nabla p_{K,h}, \boldsymbol{\psi}_{K,h}\big)_{0,K}\\ &-\frac{1}{2\rho c} \sum_{f \in \mathcal{F}_K} \big([p_h]_{K,f} + \rho c \mathbf{n}_K \cdot [\mathbf{v}_h]_{K,f}, \varphi_{K,h} + \rho c \boldsymbol{\psi}_{K,h} \cdot \mathbf{n}_K\big)_{0,f} .\end{aligned}$$

On boundary faces $f = \partial K \cap \partial\Omega$, we set $[p_h]_{K,f} = 0$ and $[\mathbf{v}_h]_{K,f} \cdot \mathbf{n}_K = -2\mathbf{v}_{K,h} \cdot \mathbf{n}_K$ for Neumann boundary conditions. This yields

$$\big(A_h(p_h, \mathbf{v}), (p_h, \mathbf{v}_h)\big)_{0,\Omega} \quad = \quad \frac{1}{2} \sum_K \sum_{f \in \mathcal{F}_K} \Big(\frac{1}{\rho c} \big\|[p_h]_{K,f}\big\|_{0,f}^2 + \rho c \big\|\mathbf{n}_K \cdot [\mathbf{v}_h]_{K,f}\big\|_{0,f}^2\Big).$$

For *electro-magnetic waves* with $(\mathbf{H}_h, \mathbf{E}_h) \in V_h$ and $(\boldsymbol{\varphi}_{K,h}, \boldsymbol{\psi}_{K,h}) \in V_{K,h}$ we have

$$\begin{aligned}\big(A_h(\mathbf{H}_h, \mathbf{E}_h), (\boldsymbol{\varphi}_{K,h}, \boldsymbol{\psi}_{K,h})\big)_{0,K} \quad = \quad &(\operatorname{curl} \mathbf{E}_{K,h}, \boldsymbol{\varphi}_{K,h})_{0,K} - \big(\operatorname{curl} \mathbf{H}_{K,h}, \boldsymbol{\psi}_{K,h}\big)_{0,K}\\ &+\frac{1}{2\sqrt{\mu\varepsilon}} \sum_{f \in \mathcal{F}_K} \Big(\big(\sqrt{\mu}[\mathbf{H}] \cdot (\mathbf{n} \times \boldsymbol{\tau}) - \sqrt{\varepsilon}[\mathbf{E}] \cdot \boldsymbol{\tau}, \sqrt{\mu}\boldsymbol{\varphi}_{K,h} \cdot (\mathbf{n} \times \boldsymbol{\tau}) - \sqrt{\varepsilon}\boldsymbol{\psi}_{K,h} \cdot \boldsymbol{\tau}\big)_{0,f}\\ &-\big(\sqrt{\mu}[\mathbf{H}] \cdot \boldsymbol{\tau} + \sqrt{\varepsilon}[\mathbf{E}] \cdot (\mathbf{n} \times \boldsymbol{\tau}), \sqrt{\mu}\boldsymbol{\varphi}_{K,h} \cdot \boldsymbol{\tau} + \sqrt{\varepsilon}\boldsymbol{\psi}_{K,h} \cdot (\mathbf{n} \times \boldsymbol{\tau})\big)_{0,f} .\end{aligned}$$

The perfect conducting boundary conditions on the faces $f = \partial K \cap \partial\Omega$ are modeled by the (only virtual) definition of $\mathbf{n}_K \times \mathbf{E}_{K_f} = -\mathbf{n}_K \times \mathbf{E}_K$ and $\mathbf{n}_K \times \mathbf{H}_{K_f} = \mathbf{n}_f \times \mathbf{H}_K$, i.e., $\mathbf{n}_K \times [\mathbf{E}]_{K,f} = -2\mathbf{n}_K \times \mathbf{E}_K$ and $\mathbf{n}_K \times [\mathbf{H}]_{K,f} = \mathbf{0}$. This yields

$$\big(A_h(\mathbf{H}_h, \mathbf{E}_h), (\mathbf{H}_h, \mathbf{E}_h)\big)_{0,\Omega} = \frac{1}{2} \sum_K \sum_{f \in \mathcal{F}_K} \left(\frac{1}{\sqrt{\varepsilon}} \big\|\mathbf{n}_K \times [\mathbf{H}_h]_{K,f}\big\|_{0,f}^2 + \frac{1}{\sqrt{\mu}} \big\|\mathbf{n}_K \times [\mathbf{E}_h]_{K,f}\big\|_{0,f}^2 \right) .$$

4 A Petrov–Galerkin space-time discretization

Let $\overline{Q} = \bigcup_{R\in\mathcal{R}} \overline{R}$ be a decomposition of the space-time cylinder into space-time cells $R = I \times K$ with $K \subset \Omega$ and $I = (t_-, t_+) \subset (0,T)$; $\mathcal{R}$ denotes the set of space-time cells. For every $R \in \mathcal{R}$ we choose local ansatz and test spaces $V_{h,R}, W_{h,R} \subset \mathrm{L}_2(R;\mathbb{R}^J)$ with $W_{h,R} \subset \partial_t V_{h,R}$, and we define the global ansatz and test space

$$\begin{aligned} V_h &= \Big\{ \mathbf{v}_h \in \mathrm{H}^1((0,T);H) \colon \mathbf{v}_h(0,\mathbf{x}) = \mathbf{0} \text{ for a.a. } \mathbf{x}\in\Omega \text{ and } \mathbf{v}_{h,R} = \mathbf{v}_h|_R \in V_{h,R} \Big\}\,, \\ W_h &= \Big\{ \mathbf{w}_h \in \mathrm{L}_2((0,T);H) \colon \mathbf{w}_{h,R} = \mathbf{w}_h|_R \in W_{h,R} \Big\}\,. \end{aligned}$$

By construction, functions in W_h are discontinuous in space and time, and functions in V_h are continuous in time, i.e., $\mathbf{v}_h(\mathbf{x},\cdot)$ is continuous on $[0,T]$ for a.a. $\mathbf{x}\in\Omega$.

In addition we aim for $\dim(V_h) = \dim(W_h)$, which restricts the choice of $V_{h,R}$. In the most simple case this can be achieved for a tensor product space-time discretization with a fixed mesh $\mathcal{K}$ in space and a time series $0 = t_0 < t_1 < \cdots < t_N = T$, i.e., $\mathcal{R} = \bigcup_{K\in\mathcal{K}} \bigcup_{n=1}^N (t_{n-1},t_n)\times K$. Then, we can select a discrete space H_h with $H_{h,K} = \mathbb{P}_p(K;\mathbb{R}^J)$ independently of t, and in every time slice we define $W_{h,R} = H_{h,K}$ constant in time on $R = (t_{n-1},t_n)\times K$. This yields in this case piecewise linear approximations in time

$$\begin{aligned} V_h = \Big\{ \mathbf{v}_h \in \mathrm{H}^1((0,T);H) \,:\, & \mathbf{v}_h(0,\mathbf{x}) = \mathbf{0}\,,\ \mathbf{v}_h(t_n,\mathbf{x}) \in H_h \text{ for a.a. } \mathbf{x}\in\Omega \text{ and } n=1,\dots,N\,, \text{ and} \\ & \mathbf{v}_h(t,\mathbf{x}) = \frac{t_n - t}{t_n - t_{n-1}}\mathbf{v}_h(t_{n-1},\mathbf{x}) + \frac{t - t_{n-1}}{t_n - t_{n-1}}\mathbf{v}_h(t_n,\mathbf{x}) \text{ for } t\in(t_{n-1},t_n) \Big\}\,. \end{aligned}$$

In the general case, we select locally in space and time polynomial degrees p_R and q_R in R, and we set for the local test space $W_{h,R} = \big(\mathbb{P}_{q_R-1}(I;\mathbb{R}^J \otimes \mathbb{P}_{p_R}(K;\mathbb{R}^J))\big)$. Then we define for $R\in\mathcal{R}$

$$\begin{aligned} V_{h,R} = \Big\{ \mathbf{v}_{h,R} \in \mathrm{L}_2(R;\mathbb{R}^J) \,:\, & \mathbf{v}_{h,R}(t,\mathbf{x}) = \frac{t_+ - t}{t_+ - t_-}\mathbf{v}_h(t_-,\mathbf{x}) + \frac{t - t_-}{t_+ - t_-}\mathbf{w}_{h,R}(t,\mathbf{x})\,, \\ & \mathbf{v}_h \in V_h|_{[0,t_-]}\,,\ \mathbf{w}_{h,R}\in W_{h,R}\,,\ (t,\mathbf{x}) \in R = (t_-,t_+)\times K \Big\}\,. \end{aligned}$$

This yields $\mathbf{v}_{h,R}(\cdot,\mathbf{x}) \in \mathbb{P}_{q_R}(I;\mathbb{R}^J)$ for $\mathbf{v}_{h,R}\in V_{h,R}$ and $(\cdot,\mathbf{x})\in R$.

The discontinuous Galerkin operator in space is extended to the space-time operator $A_h\mathbf{v}_h \in W_h$ by defining

$$\begin{aligned} (A_h\mathbf{v}_h,\mathbf{w}_h)_{0,Q} = \sum_{R=I\times K} \Big(& \big(\operatorname{div}\mathbf{F}(\mathbf{v}_{h,R}),\mathbf{w}_{h,R}\big)_{0,R} \\ & + \sum_{f\in\mathcal{F}_K} \big(\mathbf{n}_K\cdot(\mathbf{F}_K^{\mathrm{up}}(\mathbf{v}_h) - \mathbf{F}(\mathbf{v}_{h,R})),\mathbf{w}_{h,R}\big)_{0,I\times f}\Big) \end{aligned} \tag{13}$$

for $\mathbf{v}_h\in V_h$ and $\mathbf{w}_h\in W_h$. We define the discrete space-time operator $L_h\in\mathcal{L}(V_h,W_h)$ and the corresponding discrete bilinear form $b_h(\cdot,\cdot) = \big(L_h\cdot,\cdot\big)_{0,Q}$ by

$$\big(L_h\mathbf{v}_h,\mathbf{w}_h\big)_{0,Q} = \big(M_h\partial_t\mathbf{v}_h + A_h\mathbf{v}_h,\mathbf{w}_h\big)_{0,Q}\,.$$

In order to show that a solution to our Petrov–Galerkin scheme exists, we check the inf-sup stability of the discrete bilinear form $b_h(\cdot,\cdot)$ with respect to the discrete norm

$$\|\mathbf{v}_h\|_{V_h}^2 = \|\mathbf{v}_h\|_W^2 + \|M_h^{-1}L_h\mathbf{v}_h\|_W^2\,.$$

By construction, $b_h(\cdot,\cdot)$ is bounded in $V_h\times W_h$, i.e.,

$$b_h(\mathbf{v}_h,\mathbf{w}_h) = \big(L_h\mathbf{v}_h,\mathbf{w}_h\big)_{0,Q} \le \|M_h^{-1}L_h\mathbf{v}_h\|_W\|\mathbf{w}_h\|_W \le \|\mathbf{v}_h\|_{V_h}\|\mathbf{w}_h\|_W\,,\qquad \mathbf{v}_h\in V_h\,,\ \mathbf{w}_h\in W_h\,.$$

For the verification of the inf-sup stability, we introduce the L_2-projection $\Pi_h \colon W \to W_h$ which is defined by $\big(\Pi_h \mathbf{v}, \mathbf{w}_h\big)_{0,Q} = \big(\mathbf{v}, \mathbf{w}_h\big)_{0,Q}$ for $\mathbf{w}_h \in W_h$. Then, by construction, $\Pi_h A_h = A_h$ and $\Pi_h L_h = L_h$. Moreover, we define the non-negative weight function in time $d_T(t) = T - t$, and we observe

$$\int_0^T \int_0^t \phi(s)\,\mathrm{d}s\,\mathrm{d}t = \int_0^T d_T(t)\phi(t)\,\mathrm{d}t\,, \qquad \phi \in \mathrm{L}_1(0,T)\,.$$

Lemma 4.1 *Assume that*

$$\big(M_h \partial_t \mathbf{v}_h, d_T \mathbf{v}_h\big)_{0,Q} \le \big(L_h \mathbf{v}_h, d_T \Pi_h \mathbf{v}_h\big)_{0,Q}\,, \qquad \mathbf{v}_h \in V_h\,. \tag{14}$$

Then, the bilinear form $b_h(\cdot,\cdot)$ is inf-sup stable in $V_h \times W_h$ with $\beta = 1/\sqrt{1+4T^2}$, i.e.,

$$\sup_{\mathbf{w}_h \in W_h} \frac{b_h(\mathbf{v}_h, \mathbf{w}_h)}{\|\mathbf{w}_h\|_W} \ge \beta\, \|\mathbf{v}_h\|_{V_h}\,, \qquad \mathbf{v}_h \in V_h\,.$$

Proof. Transferring the proof of Lem. 2.1 to the discrete setting yields

$$\begin{aligned}
\|\mathbf{v}_h\|_W^2 &= \int_0^T \big(M_h \mathbf{v}_h(t), \mathbf{v}_h(t)\big)_{0,\Omega}\,\mathrm{d}t = \int_0^T \Big(\big(M_h \mathbf{v}_h(t), \mathbf{v}_h(t)\big)_{0,\Omega} - \big(M_h \mathbf{v}_h(0), \mathbf{v}_h(0)\big)_{0,\Omega}\Big)\,\mathrm{d}t\\
&= \int_0^T \int_0^t \partial_t \big(M_h \mathbf{v}_h(s), \mathbf{v}_h(s)\big)_{0,\Omega}\,\mathrm{d}s\,\mathrm{d}t = 2\int_0^T \int_0^t \big(M_h \partial_t \mathbf{v}_h(s), \mathbf{v}_h(s)\big)_{0,\Omega}\,\mathrm{d}s\,\mathrm{d}t\\
&= 2\big(M_h \partial_t \mathbf{v}_h, d_T \mathbf{v}_h\big)_{0,Q} \le 2\big(L_h \mathbf{v}_h, d_T \Pi_h \mathbf{v}_h\big)_{0,Q} \le 2T\, \|M_h^{-1} L_h \mathbf{v}_h\|_W \|\mathbf{v}_h\|_W\,.
\end{aligned}$$

This yields $\|\mathbf{v}_h\|_W \le 2T\,\|M_h^{-1} L_h \mathbf{v}_h\|_W$ and thus $\|\mathbf{v}_h\|_{V_h} \le \sqrt{1+4T^2}\,\|M_h^{-1} L_h \mathbf{v}_h\|_W$, which implies the inf-sup stability using $b_h(\mathbf{v}_h, \mathbf{w}_h) = (L_h \mathbf{v}_h, \mathbf{w}_h)_{0,Q} = (M_h^{-1} L_h \mathbf{v}_h, \mathbf{w}_h)_W$ and inserting $\mathbf{w}_h = M_h^{-1} L_h \mathbf{v}_h$

$$\sup_{\mathbf{w}_h \in W_h \setminus \{\mathbf{0}\}} \frac{b_h(\mathbf{v}_h, \mathbf{w}_h)}{\|\mathbf{w}_h\|_W} = \sup_{\mathbf{w}_h \in W_h \setminus \{\mathbf{0}\}} \frac{(M_h^{-1} L_h \mathbf{v}_h, \mathbf{w}_h)_W}{\|\mathbf{w}_h\|_W} \ge \|M_h^{-1} L_h \mathbf{v}_h\|_W \ge \frac{1}{\sqrt{1+4T^2}} \|\mathbf{v}_h\|_{V_h}\,.$$

□

As in Thm. 2.2, this shows the existence of a unique discrete Petrov–Galerkin solution (provided that the assumption in Lem. 4.1 is satisfied).

Theorem 4.2 *For given $\mathbf{f} \in \mathrm{L}_2(Q;\mathbb{R}^J)$ there exists a unique solution $\mathbf{u}_h \in V_h$ of*

$$(L_h \mathbf{u}_h, \mathbf{w}_h)_{0,Q} = (\mathbf{f}, \mathbf{w}_h)_{0,Q}\,, \qquad \mathbf{w}_h \in W_h\,, \tag{15}$$

satisfying the a priori bound $\|\mathbf{u}_h\|_{V_h} \le \sqrt{4T^2+1}\|M_h^{-1}\Pi_h \mathbf{f}\|_W$.

The convergence will be analyzed with respect to the discrete norm $\|\cdot\|_{V_h}$. For $\mathbf{v} \in V$ the consistency of the numerical flux in (13) yields $(A_h \mathbf{v}, \mathbf{w}_h)_{0,Q} = \big(\operatorname{div} \mathbf{F}(\mathbf{v}), \mathbf{w}_h\big)_{0,Q}$ so that $A_h \mathbf{v} = \Pi_h \operatorname{div} \mathbf{F}(\mathbf{v})$. This shows that A_h and thus also $\|\cdot\|_{V_h}$ can be evaluated in $V + V_h$ and that $b_h(\cdot,\cdot)$ is continuous with respect to this extension.

Theorem 4.3 *Let $\mathbf{u} \in V$ be the solution of (1) and $\mathbf{u}_h \in V_h$ its approximation solving (15). Then, we have*

$$\|\mathbf{u} - \mathbf{u}_h\|_{V_h} \le (1 + \beta^{-1}) \inf_{\mathbf{v}_h \in V_h} \|\mathbf{u} - \mathbf{v}_h\|_{V_h}\,.$$

If in addition the solution is sufficiently smooth, we obtain the a priori error estimate

$$\|\mathbf{u}-\mathbf{u}_h\|_{V_h} \le C\big(\triangle t^q + \triangle x^p\big)\Big(\|\partial_t^{q+1}\mathbf{u}\|_{0,Q} + \|\mathrm{D}^{p+1}\mathbf{u}\|_{0,Q}\Big)$$

for $\triangle t$, $\triangle x$ *and* $p, q \ge 1$ *with* $\triangle t \ge t_+ - t_-$, $\triangle x \ge \operatorname{diam}(K)$, $p \le p_R$ *and* $q \le q_R$ *for all* $R = (t_-, t_+) \times K$.

Proof. The consistency (10) of the discontinuous Galerkin method yields $(A_h\mathbf{u}(t), \mathbf{w}_h(t))_{0,\Omega} = (A\mathbf{u}(t), \mathbf{w}_h(t))_{0,\Omega}$ and thus also consistency of the Petrov–Galerkin setting, i.e., $b_h(\mathbf{u}, \mathbf{w}_h) = b(\mathbf{u}, \mathbf{w}_h) = (\mathbf{f}, \mathbf{w}_h)_{0,Q} = b_h(\mathbf{u}_h, \mathbf{w}_h)$. This gives for all $\mathbf{v}_h \in V_h$ and $\mathbf{w}_h \in W_h$

$$b_h(\mathbf{v}_h - \mathbf{u}_h, \mathbf{w}_h) = b_h(\mathbf{v}_h - \mathbf{u}, \mathbf{w}_h) \le \|\mathbf{v}_h - \mathbf{u}\|_{V_h}\|\mathbf{w}_h\|_W$$

and thus

$$\begin{aligned}\|\mathbf{u}-\mathbf{u}_h\|_{V_h} &\le \|\mathbf{u}-\mathbf{v}_h\|_{V_h} + \|\mathbf{v}_h-\mathbf{u}_h\|_{V_h}\\ &\le \|\mathbf{u}-\mathbf{v}_h\|_{V_h} + \beta^{-1}\sup_{\mathbf{w}_h\in W_h\setminus\{\mathbf{0}\}}\frac{b_h(\mathbf{v}_h-\mathbf{u}_h,\mathbf{w}_h)}{\|\mathbf{w}_h\|_W} \le (1+\beta^{-1})\,\|\mathbf{u}-\mathbf{v}_h\|_{V_h}\,.\end{aligned}$$

Now we assume that the solution is regular satisfying $\mathbf{u} \in \mathrm{H}^{q+1}\big((0,T);\mathrm{L}_2(\Omega;\mathbb{R}^J)\big) \cap \mathrm{L}_2\big((0,T);\mathrm{H}^{p+1}(\Omega;\mathbb{R}^J)\big)$. We have by consistency $A_h\mathbf{v}_h = A\mathbf{v}_h$ for all $\mathbf{v}_h \in V_h \cap \mathrm{H}^1(\Omega;\mathbb{R}^J)$, so that the error estimate yields

$$\|\mathbf{u}-\mathbf{u}_h\|_{V_h} \le (1+\beta^{-1})\inf_{\mathbf{v}_h\in V_h\cap\mathrm{H}^1(\Omega;\mathbb{R}^J)}\|\mathbf{u}-\mathbf{v}_h\|_{V_h} \le C\Big(\|\partial_t(\mathbf{u}-I_h\mathbf{u})\|_{0,Q} + \|\mathrm{D}(\mathbf{u}-I_h\mathbf{u})\|_{0,Q}\Big)\,,$$

where $I_h\colon V \to V_h \cap \mathrm{H}^1(\Omega;\mathbb{R}^J)$ is a suitable Clément-type interpolation operator. By standard assumptions on the right-hand side and the mesh regularity we obtain a bound depending on $\triangle t$ in time and $\triangle x$ in space. □

We check the assumptions of Lem. 4.1 for the special case of a tensor product discretization, where the polynomial degrees in time are fixed on every time slice $I = (t_{n-1}, t_n) \subset (0,T)$ and the polynomial degrees in space are fixed on every $K \subset \Omega$. Then we have the local spaces $W_{h,R} = \mathbb{P}_{p_K}(K) \otimes \mathbb{P}_{q_I-1}$ and $V_R = \mathbb{P}_{p_K}(K) \otimes \mathbb{P}_{q_I}$ on a space-time cell $R = I \times K \in \mathcal{R}$, i.e., $p_R = p_K$ and $q_R = q_I$. Note that for $q_I \equiv 1$ the Petrov–Galerkin method in time is equivalent to the implicit midpoint rule, see also [BR99].

Lemma 4.4 *In the case of tensor product space-time discretizations, the condition (14) in Lem. 4.1 is satisfied: we have for $\mathbf{v}_h \in V_h$*

$$\Pi_h \partial_t \mathbf{v}_h = \partial_t \mathbf{v}_h\,, \qquad \big(M_h \partial_t \mathbf{v}_h, d_T \mathbf{v}_h\big)_{0,Q} \le \big(M_h \partial_t \mathbf{v}_h, d_T \Pi_h \mathbf{v}_h\big)_{0,Q}\,, \qquad 0 \le \big(A_h \mathbf{v}_h, d_T \Pi_h \mathbf{v}_h\big)_{0,Q}\,.$$

Proof. Let H_h be the discontinuous Galerkin space in Ω with $H_{h,K} = \mathbb{P}_{p_K}(K)$. In the tensor product case, for $\mathbf{v}_h \in V_h$ and $\mathbf{w}_h \in W_h$ representations exist in the form

$$\mathbf{v}_h(t,\mathbf{x}) = \sum_{k=0}^{q_I} \boldsymbol{\psi}_{I,k}(\mathbf{x})\lambda_{I,k}(t)\,, \qquad \mathbf{w}_h(t,\mathbf{x}) = \sum_{k=0}^{q_I-1} \boldsymbol{\phi}_{I,k}(\mathbf{x})\lambda_{I,k}(t)\,, \qquad (t,\mathbf{x}) \in \bigcup I \times K$$

with orthonormal Legendre polynomials $\lambda_{I,k} \in \mathbb{P}_k$ in $\mathrm{L}_2(I)$ and $\boldsymbol{\psi}_{I,k}, \boldsymbol{\phi}_{I,k} \in H_h$. We observe

$$\partial_t \mathbf{v}_{h,R}(t,\mathbf{x}) = \sum_{k=1}^{q_I} \boldsymbol{\psi}_{I\times K,k}(\mathbf{x})\partial_t\lambda_{I,k}(t)\,, \qquad \partial_t \lambda_{I,k} \in \mathbb{P}_{k-1}\,, \quad (t,\mathbf{x}) \in R = I \times K\,,$$

i.e., $\partial_t \mathbf{v}_{h,R} \in W_{h,R}$ and thus $\Pi_h \partial_t \mathbf{v}_h = \partial_t \mathbf{v}_h$. Furthermore, we have

$$\begin{aligned}
\big(d_T M_h \partial_t \mathbf{v}_h, \mathbf{v}_h - \Pi_h \mathbf{v}_h\big)_{0,Q} &= \sum_I \sum_{k=0}^{q_I} \big(M_h \boldsymbol{\psi}_{I,k}, \boldsymbol{\psi}_{I,q_I}\big)_{0,\Omega} \big(d_T \partial_t \lambda_{I,k}, \lambda_{I,q_I}\big)_{0,I} \\
&= \sum_I \big(M_h \boldsymbol{\psi}_{I,q_I}, \boldsymbol{\psi}_{I,q_I}\big)_{0,\Omega} \big(d_T \partial_t \lambda_{I,q_I}, \lambda_{I,q_I}\big)_{0,I} = -k \sum_I \big(M_h \boldsymbol{\psi}_{I,q_I}, \boldsymbol{\psi}_{I,q_I}\big)_{0,\Omega} \le 0
\end{aligned}$$

since $\big(d_T \partial_t \lambda_{I,k}, \lambda_{I,q_I}\big)_{0,I} = 0$ for $k < q_I$ and $\big(d_T \partial_t \lambda_{I,q_I}, \lambda_{I,q_I}\big)_{0,I} = -\big(t\partial_t \lambda_{I,q_I}, \lambda_{I,q_I}\big)_{0,I} = -q_I$ (see Lem. 7.1 in the appendix for a proof). From

$$\begin{aligned}
\big(A_h \mathbf{v}_h, \mathbf{w}_h\big)_{0,Q} &= \sum_I \sum_K \Big(\big(\operatorname{div} \mathbf{F}(\mathbf{v}_{h,R}), \mathbf{w}_{h,R}\big)_{0,I\times K} + \sum_{f \in \mathcal{F}_K} \big(\mathbf{n}_K \cdot (\mathbf{F}_K^{\mathrm{up}}(\mathbf{v}_h) - \mathbf{F}(\mathbf{v}_{h,R})), \mathbf{w}_{h,R}\big)_{0,I\times f}\Big) \\
&= \sum_I \sum_K \sum_{k=0}^{q_I} \sum_{j=0}^{q_I-1} \Big(\big(\operatorname{div} \mathbf{F}(\boldsymbol{\psi}_{I\times K,k}), \boldsymbol{\phi}_{I\times K,j}\big)_{0,K} \\
&\qquad + \sum_{f \in \mathcal{F}_K} \big(\mathbf{n}_K \cdot (\mathbf{F}_K^{\mathrm{up}}(\boldsymbol{\psi}_{I,k}) - \mathbf{F}(\boldsymbol{\psi}_{I\times K,k})), \boldsymbol{\phi}_{I\times K,j}\big)_{0,f}\Big) \big(\lambda_{I,k}, \lambda_{I,j}\big)_{0,I} \\
&= \big(A_h \Pi_h \mathbf{v}_h, \mathbf{w}_h\big)_{0,Q}
\end{aligned}$$

we obtain in the tensor product case $A_h = A_h \Pi_h$ and thus

$$\begin{aligned}
\big(A_h \mathbf{v}_h, d_T \Pi_h \mathbf{v}_h\big)_{0,Q} &= \big(A_h \Pi_h \mathbf{v}_h, d_T \Pi_h \mathbf{v}_h\big)_{0,Q} \\
&= \sum_I \sum_K \sum_{k=0}^{q_I-1} \sum_{j=0}^{q_I-1} \Big(\big(\operatorname{div} \mathbf{F}(\boldsymbol{\psi}_{I\times K,k}), \boldsymbol{\psi}_{I\times K,j}\big)_{0,K} \\
&\qquad + \sum_{f \in \mathcal{F}_K} \big(\mathbf{n}_K \cdot (\mathbf{F}_K^{\mathrm{up}}(\boldsymbol{\psi}_{I,k}) - \mathbf{F}(\boldsymbol{\psi}_{I\times K,k})), \boldsymbol{\psi}_{I\times K,j}\big)_{0,f}\Big) \big(\lambda_{I,k}, d_T\lambda_{I,j}\big)_{0,I} \\
&= \sum_I \sum_{k=0}^{q_I-1} \sum_{j=0}^{q_I-1} \big(A_h \boldsymbol{\psi}_{I,k}, \boldsymbol{\psi}_{I,j}\big)_{0,\Omega} (\lambda_{I,k}, d_T \lambda_{I,j})_{0,I} \ge 0
\end{aligned}$$

since both matrices with entries $\big(A_h \boldsymbol{\psi}_{I,k}, \boldsymbol{\psi}_{I,j}\big)_{0,\Omega}$ and $(\lambda_{I,k}, d_T \lambda_{I,j})_{0,I}$, respectively, are positive semi-definite. $\square$

5 Duality based goal-oriented error estimation

In order to develop an adaptive strategy for the selection of the local polynomial degrees p_R, q_R we derive an error indicator with respect to a given linear goal functional $E \in W'$. Following the framework in [BR03], we define the adjoint problem and solve the dual problem. Then, the error is estimated in terms of the local residual and the dual weight.

The adjoint operator L^* in space and time is defined on the adjoint Hilbert space

$$V^* = \big\{\mathbf{w} \in W \colon \text{there exists } \mathbf{g} \in W \text{ such that } (L\mathbf{v},\mathbf{w})_{0,Q} = (\mathbf{v},\mathbf{g})_{0,Q} \text{ for all } \mathbf{v} \in V\big\}$$

and is characterized by

$$(\mathbf{v}, L^*\mathbf{w})_{0,Q} = (L\mathbf{v},\mathbf{w})_{0,Q}\,, \qquad \mathbf{v} \in V\,,\ \mathbf{w} \in V^*\,.$$

Then, V^* is the closure of $\big\{\mathbf{v}^* \in \mathrm{C}^1([0,T];\mathcal{D}(A^*))\colon \mathbf{v}^*(T) = \mathbf{0}\big\}$ and $L^* = -L$ on $V \cap V^*$.

For the evaluation of the error functional E we introduce the dual solution $\mathbf{u}^* \in V^*$ with

$$(\mathbf{w}, L^*\mathbf{u}^*)_{0,Q} = \langle E, \mathbf{w}\rangle\,, \qquad \mathbf{w} \in W\,.$$

Let $\mathbf{u} \in V$ be the solution of (1), and $\mathbf{u}_h \in V_h$ its approximation solving (15). Now we derive an exact error representation for the error functional in the case that the dual solution is sufficiently smooth such that $\mathbf{u}^*(t,\cdot)|_f \in \mathrm{L}_2(f;\mathbb{R}^J)$ for all faces $f \in \mathcal{F}_h$ and a.a. $t \in (0,T)$. Inserting the consistency of the numerical flux (10) yields for all $\mathbf{w}_h \in W_h$

$$\begin{aligned}
\langle E, \mathbf{u}-\mathbf{u}_h\rangle &= \big(\mathbf{u}-\mathbf{u}_h, -M\partial_t\mathbf{u}^* - \operatorname{div}\mathbf{F}(\mathbf{u}^*)\big)_{0,Q}\\
&= \big(\mathbf{u}, -M\partial_t\mathbf{u}^* - \operatorname{div}\mathbf{F}(\mathbf{u}^*)\big)_{0,Q} - \big(\mathbf{u}_h, -M\partial_t\mathbf{u}^* - \operatorname{div}\mathbf{F}(\mathbf{u}^*)\big)_{0,Q}\\
&= \big(M\partial_t\mathbf{u} + \operatorname{div}\mathbf{F}(\mathbf{u}), \mathbf{u}^*\big)_{0,Q} - \big(\mathbf{u}, \mathbf{n}\cdot\mathbf{F}(\mathbf{u}^*)\big)_{0,\partial Q}\\
&\qquad - \sum_{R\in\mathcal{R}} \Big(\big(M\partial_t\mathbf{u}_h + \operatorname{div}\mathbf{F}(\mathbf{u}_h), \mathbf{u}^*\big)_{0,R} - \big(\mathbf{u}_h, \mathbf{n}_R\cdot\mathbf{F}(\mathbf{u}^*)\big)_{0,\partial R}\Big)\\
&= \big(\mathbf{f}, \mathbf{u}^*\big)_{0,Q} - \sum_{R=I\times K\in\mathcal{R}} \Big(\big(M\partial_t\mathbf{u}_h + \operatorname{div}\mathbf{F}(\mathbf{u}_h), \mathbf{u}^*\big)_{0,R} - \big(\mathbf{u}_h, \mathbf{n}_K\cdot\mathbf{F}(\mathbf{u}^*)\big)_{0,I\times\partial K}\Big)\\
&= \sum_{R=I\times K\in\mathcal{R}} \Big(\big(\mathbf{f} - M\partial_t\mathbf{u}_h - \operatorname{div}\mathbf{F}(\mathbf{u}_h), \mathbf{u}^*\big)_{0,R} + \big(\mathbf{n}_K\cdot\mathbf{F}(\mathbf{u}_h), \mathbf{u}^*\big)_{0,I\times\partial K}\Big)\\
&= \sum_{R=I\times K\in\mathcal{R}} \Big(\big(\mathbf{f} - M\partial_t\mathbf{u}_h - \operatorname{div}\mathbf{F}(\mathbf{u}_h), \mathbf{u}^*\big)_{0,R} + \big(\mathbf{n}_K\cdot(\mathbf{F}(\mathbf{u}_h) - \mathbf{F}_K^{\mathrm{up}}(\mathbf{u}_h)), \mathbf{u}^*\big)_{0,I\times\partial K}\Big)\\
&= \sum_{R=I\times K\in\mathcal{R}} \Big(\big(\mathbf{f} - M\partial_t\mathbf{u}_h - \operatorname{div}\mathbf{F}(\mathbf{u}_h), \mathbf{u}^*-\mathbf{w}_h\big)_{0,R} + \big(\mathbf{n}_K\cdot(\mathbf{F}(\mathbf{u}_h) - \mathbf{F}_K^{\mathrm{up}}(\mathbf{u}_h)), \mathbf{u}^*-\mathbf{w}_h\big)_{0,I\times\partial K}\Big)\,.
\end{aligned}$$

From this error representation, inserting some projection $\mathbf{w}_h = \Pi_h\mathbf{u}^*$, we obtain the estimate

$$\begin{aligned}
\big|\langle E, \mathbf{u}-\mathbf{u}_h\rangle\big| \le \sum_{R=I\times K\in\mathcal{R}} \Big(&\big\|M\partial_t\mathbf{u}_h + \operatorname{div}\mathbf{F}(\mathbf{u}_h) - \mathbf{f}\big\|_{0,R}\big\|\mathbf{u}^* - \Pi_h\mathbf{u}^*\|_{0,R}\\
&+ \big\|\mathbf{n}_K\cdot(\mathbf{F}(\mathbf{u}_h) - \mathbf{F}_K^{\mathrm{up}}(\mathbf{u}_h))\big\|_{0,I\times\partial K}\big\|\mathbf{u}^* - \Pi_h\mathbf{u}^*\big\|_{0,I\times\partial K}\Big)\,.
\end{aligned} \tag{16}$$

However, this bound cannot be used since it depends on the unknown function $\mathbf{u}^*$. In applications, the following heuristic error bound is used instead. Let $\mathbf{u}_h^* \in W_h$ be a numerical approximation of the dual solution given by

$$b_h(\mathbf{v}_h, \mathbf{u}_h^*) = \langle E, \mathbf{v}_h\rangle\,, \qquad \mathbf{v}_h \in V_h$$

(using the transposed finite element matrix). Then we replace the projection error $\mathbf{u}^* - \Pi_h \mathbf{u}^*$ by $I_h \mathbf{u}_h^* - \mathbf{u}_h^*$, where I_h is a higher-order recovery operator (or a lower order interpolation operator). Then, the right-hand side of the error bound (16) is replaced by $\sum_{R\in\mathcal{R}} \eta_R$ with

$$\begin{aligned}\eta_R = \|\mathbf{f} - M\partial_t \mathbf{u}_h - \operatorname{div} \mathbf{F}(\mathbf{u}_h)\|_{0,R} \|I_h \mathbf{u}_h^* - \mathbf{u}_h^*\|_{0,R} \\ + \|\mathbf{n}_K \cdot (\mathbf{F}(\mathbf{u}_h) - \mathbf{F}_K^{\mathrm{up}}(\mathbf{u}_h))\|_{0,I\times\partial K} \|I_h \mathbf{u}_h^* - \mathbf{u}_h^*\|_{0,I\times\partial K}\,.\end{aligned}$$

These terms contain the given data functions $\mathbf{f}$ and M and are computed by a quadrature formula. Alternatively a term $\|\mathbf{f} - \mathbf{f}_h - (M - M_h)\partial_t \mathbf{u}_h\|_{0,R}$ could be separated to control this data error. Usually, this error contribution is of minor importance. This is especially the case in our numerical examples.

Remark 5.1 *The error indicator construction extends to nonlinear goal functionals $E \in \mathrm{C}^2(W)$. Then, the dual solution $\mathbf{u}^* \in V^*$ depends on the primal solution, i.e.,*

$$(\mathbf{w}, L^* \mathbf{u}^*)_{0,Q} = \langle E'(\mathbf{u}), \mathbf{w}\rangle\,, \quad \mathbf{w} \in W\,.$$

The estimate (16) applies also to $\big|E(\mathbf{u}) - E(\mathbf{u}_h)\big|$, since we have [HR03]

$$E(\mathbf{u}) - E(\mathbf{u}_h) = \langle E'(\mathbf{u}_h), \mathbf{u} - \mathbf{u}_h\rangle + \int_0^1 (1-s) E''(\mathbf{u}_h + s(\mathbf{u} - \mathbf{u}_h))\big[\mathbf{u} - \mathbf{u}_h, \mathbf{u} - \mathbf{u}_h\big]\,\mathrm{d}s$$

and the second term is quadratic in $\|\mathbf{u} - \mathbf{u}_h\|_{0,Q}$ and will thus be neglected. In our numerical examples E'' is constant.

In our examples we use the adaptive strategy for p-refinement described in Alg. 1. It depends on a parameter $\vartheta < 1$ for the adaptive selection criterion.

Algorithm 1 Adaptive algorithm.

1: choose low order polynomial degrees on the initial mesh
2: **while** $\max_R(p_R) < p_{\max}$ and $\max_R(q_R) < q_{\max}$ **do**
3: compute $\mathbf{u}_h$
4: compute $\mathbf{u}_h^*$ and a recovery $I_h \mathbf{u}_h^*$
5: compute η_R on every cell R
6: if the error is small enough STOP
7: mark space-time cell R if $\eta_R > \vartheta \max_{R'} \eta_{R'}$
8: increase polynomial degrees on marked cells
9: redistribute cells on processes for better load balancing

Remark 5.2 *For* acoustic waves *with $\mathcal{D}(A) = \mathrm{H}^1(\Omega) \times \mathrm{H}_0(\operatorname{div}, \Omega)$ we have $\mathbf{n}\cdot\mathbf{v}^* = 0$ on $\partial\Omega$ thus $(\mathbf{n}\cdot\mathbf{F}^{\mathrm{up}}(\mathbf{u}), \mathbf{u}^*)_{0,\partial\Omega} = 0$. Moreover,*

$$\begin{aligned}\langle E, (p - p_h, \mathbf{v} - \mathbf{v}_h)\rangle = \sum_{R=I\times K\in\mathcal{R}} \Big(\big(\mathbf{b} - \rho\partial_t \mathbf{v}_h + \nabla p_h, \mathbf{v}^*\big)_{0,R} + \big(-\rho\partial_t p_h + \operatorname{div} \mathbf{v}_h, p^*\big)_{0,R} \\ + \frac{1}{2\rho c} \sum_{f\in\mathcal{F}_K} \big([p_h]_{K,f} + \rho c \mathbf{n}_K \cdot [\mathbf{v}_h]_{K,f}, p^* + \rho c \mathbf{v}^* \cdot \mathbf{n}_K\big)_{0,f}\Big)\end{aligned}$$

shows that modified error indicators can be considered, e.g.,

$$\begin{aligned}\eta_R = \big\|\big(\mathbf{b} - \rho\partial_t \mathbf{v}_h + \nabla p_h, -\rho\partial_t p_h + \operatorname{div} \mathbf{v}_h, p^*\big)\big\|_{0,R} \|I_h(p_h^*, \mathbf{v}_h^*) - (p_h^*, \mathbf{v}_h^*)\|_{0,R} \\ + \|[p_h]_{K,f}\|_{0,I\times\partial K} \|\mathbf{n}_K \cdot (I_h \mathbf{v}_h^* - \mathbf{v}_h^*)\|_{0,I\times\partial K} + \|\mathbf{n}_K \cdot [\mathbf{v}_h]_{K,f}\|_{0,I\times\partial K} \|I_h p_h^* - p_h^*\|_{0,I\times\partial K}\,.\end{aligned}$$

6 Space-time multilevel preconditioner

In this section we address the numerical aspects in particular solution methods for the discrete hyperbolic space-time problem. First we describe the realization of our discretization using nodal basis functions in space and time, and then a multilevel preconditioner is introduced, and it is tested for different settings to derive a suitable solution strategy.

Nodal Discretization Here we consider the case of a tensor product space-time mesh $\mathcal{R} = \bigcup_{n=1}^{N} \mathcal{R}^n$ with time slices $\mathcal{R}^n = \bigcup_{K\in\mathcal{K}} (t_{n-1}, t_n) \times K$ and variable polynomial degrees p_R, q_R in every space-time cell R. Let $\{\boldsymbol{\psi}^n_{R,j}\}_{j=1,\dots,\dim W_{h,R}}$ be a basis of $W_{h,R}$ and define $W_h^n = \operatorname{span}\Big\{ \bigcup_{R\in\mathcal{R}^n} \bigcup_{j=1}^{\dim W_{h,R}} \boldsymbol{\psi}^n_{R,j} \Big\}$. Then, the solution $\mathbf{u}_h \in V_h$ is represented by

$$\mathbf{u}_h(t,\mathbf{x}) = \frac{t_n - t}{t_n - t_{n-1}} \mathbf{u}_h^{n-1}(t_{n-1},\mathbf{x}) + \frac{t - t_{n-1}}{t_n - t_{n-1}} \mathbf{u}_h^n(t,\mathbf{x}) \quad \text{for} \quad (t,\mathbf{x}) \in (t_{n-1}, t_n) \times K$$

with $\mathbf{u}_h^0 = \mathbf{0}$ and $\mathbf{u}_h^n \in W_h^n$, $n = 1,\dots,N$. The corresponding coefficient vector of the solution is denoted by $\underline{u} = (\underline{u}^1,\dots,\underline{u}^N)^\top$, where $\underline{u}^n \in \mathbb{R}^{\dim W_h^n}$ is the coefficient vector of $\mathbf{u}_h^n = \sum_{R\in\mathcal{R}^n} \sum_{j=1}^{\dim W_{h,R}} \underline{u}_{R,j} \boldsymbol{\psi}^n_{R,j}$. With respect to this basis, the discrete space-time system (1) has the matrix representation $\underline{L}\,\underline{u} = \underline{f}$ with the block matrix

$$\underline{L} = \begin{pmatrix} \underline{D}^1 & & & \\ \underline{C}^1 & \underline{D}^2 & & \\ & \ddots & \ddots & \\ & & \underline{C}^{N-1} & \underline{D}^N \end{pmatrix}$$

with matrix entries

$$\underline{D}^n_{R',k,R,j} = \int_{t_{n-1}}^{t_n} \int_\Omega L_h\Big(\frac{t - t_{n-1}}{t_n - t_{n-1}} \boldsymbol{\psi}^n_{R,j}(t,\mathbf{x})\Big) \boldsymbol{\psi}^n_{R',k}(t,\mathbf{x}) \,\mathrm{d}\mathbf{x}\mathrm{d}t\,, \qquad R, R' \in \mathcal{R}^n$$

$$\underline{C}^n_{R',k,R,j} = \int_{t_{n-1}}^{t_n} \int_\Omega L_h\Big(\frac{t_n - t}{t_n - t_{n-1}} \boldsymbol{\psi}^{n-1}_{R,j}(t_{n-1},\mathbf{x})\Big) \boldsymbol{\psi}^n_{R',k}(t,\mathbf{x}) \,\mathrm{d}\mathbf{x}\mathrm{d}t\,, \qquad R \in \mathcal{R}^{n-1},\ R' \in \mathcal{R}^n$$

and the right-hand side $\underline{f} = (\underline{f}^1,\dots,\underline{u}^N)$ with $\underline{f}^n_{j,R} = (\mathbf{f}, \boldsymbol{\psi}^n_{R,j})_{0,R}$. Sequentially, this system can be solved by a block-Gauss–Seidel method (corresponding to implicit time integration)

$$\underline{D}^1\underline{u}^1 = \underline{f}^1\,, \qquad \underline{D}^2\underline{u}^2 = \underline{f}^2 - \underline{C}^1\underline{u}^1\,, \dots, \qquad \underline{D}^N\underline{u}^N = \underline{f}^N - \underline{C}^{N-1}\underline{u}^{N-1}\,,$$

provided that $\underline{D}^n$ can be inverted efficiently.

Multilevel methods For space-time multilevel preconditioners we consider hierarchies in space and time. Therefore, let $\mathcal{R}_{0,0}$ be the coarse space-time mesh, and let $\mathcal{R}_{l,k}$ be the discretization obtained by $l = 1,\dots,l_{\text{max}}$ uniform refinements in space and $k = 1,\dots,k_{\text{max}}$ refinements in time. Let $V_{l,k}$ be the approximation spaces on $\mathcal{R}_{l,k}$ with fixed polynomial degrees $p_R \equiv p$ and $q_R \equiv q$. Let $\underline{L}_{l,k}$ be the corresponding matrix representations of the discrete operator L_h in $V_{l,k}$.

The multilevel preconditioner combines smoothing operations on different levels and requires transfer matrices between the levels. Since the spaces are nested, we can define prolongation matrices $\underline{P}^{l,k}_{l-1,k}$ and $\underline{P}^{l,k}_{l,k-1}$ representing the natural injections $V_{l-1,k} \subset V_{l,k}$ in space and $V_{l,k-1} \subset V_{l,k}$ in time. Correspondingly, the restriction matrices $\underline{R}^{l,k}_{l-1,k}$ and $\underline{R}^{l,k}_{l,k-1}$ represent the L_2-projections of the test spaces $W_{l,k} \supset W_{l-1,k}$ and $W_{l,k} \supset W_{l,k-1}$.

For the smoothing operations on level (l, k) we consider the block-Jacobi preconditioner or the block-Gauss–Seidel preconditioner (where all components corresponding to a space-time cell R build blocks)

$$\underline{B}^{\mathsf{J}}_{l,k} = \theta_{l,k}\, \text{block\_diag}(\underline{L}_{l,k})^{-1}\,, \qquad \underline{B}^{\mathsf{GS}}_{l,k} = \theta_{l,k} \left(\text{block\_lower}(\underline{L}_{l,k}) + \text{block\_diag}(\underline{L}_{l,k})\right)^{-1}$$

with damping parameter $\theta_{l,k} \in (0, 1]$. The corresponding iteration matrices are given by $\underline{S}^{\mathsf{J}}_{l,k} = \mathrm{id}_{l,k} - \underline{B}^{\mathsf{J}}_{l,k}\underline{L}_{l,k}$ and $\underline{S}^{\mathsf{GS}}_{l,k} = \mathrm{id}_{l,k} - \underline{B}^{\mathsf{GS}}_{l,k}\underline{L}_{l,k}$, and the number of pre- and postsmoothing steps are denoted by $\nu^{\mathsf{pre}}_{l,k}$ and $\nu^{\mathsf{post}}_{l,k}$.

Now, the multilevel preconditioner $\underline{B}^{\mathsf{ML}}_{l,k}$ is defined recursively. On the coarse level, we use a parallel direct linear solver $\underline{B}^{\mathsf{ML}}_{0,0} = (\underline{L}_{0,0})^{-1}$. Then, we have two options: restricting in time defines $\underline{B}^{\mathsf{ML}}_{l,k}$ by

$$\mathrm{id}_{l,k} - \underline{B}^{\mathsf{ML}}_{l,k}\underline{L}_{l,k} = \left(\mathrm{id}_{l,k} - \underline{B}^{\mathsf{J}}_{l,k}\underline{L}_{l,k}\right)^{\nu^{\mathsf{pre}}_{l,k}} \left(\mathrm{id}_{l,k} - \underline{P}^{l,k}_{l,k-1}\underline{B}^{\mathsf{ML}}_{l,k-1}\underline{R}^{l,k}_{l,k-1}\underline{L}_{l,k}\right) \left(\mathrm{id}_{l,k} - \underline{B}^{\mathsf{J}}_{l,k}\underline{L}_{l,k}\right)^{\nu^{\mathsf{post}}_{l,k}}$$

with Jacobi smoothing (cf. Fig. 2), and restricting in space yields

$$\mathrm{id}_{l,k} - \underline{B}^{\mathsf{ML}}_{l,k}\underline{L}_{l,k} = \left(\mathrm{id}_{l,k} - \underline{B}^{\mathsf{GS}}_{l,k}\underline{L}_{l,k}\right)^{\nu^{\mathsf{pre}}_{l,k}} \left(\mathrm{id}_{l,k} - \underline{P}^{l,k}_{l,k-1}\underline{B}^{\mathsf{ML}}_{l-1,k}\underline{R}^{l,k}_{l,k-1}\underline{L}_{l,k}\right) \left(\mathrm{id}_{l,k} - \underline{B}^{\mathsf{GS}}_{l,k}\underline{L}_{l,k}\right)^{\nu^{\mathsf{post}}_{l,k}}$$

with Gauss–Seidel smoothing, cf. Fig. 3 for an illustration of the two options and Alg. 2 for the recursive realization of the multilevel preconditioner.

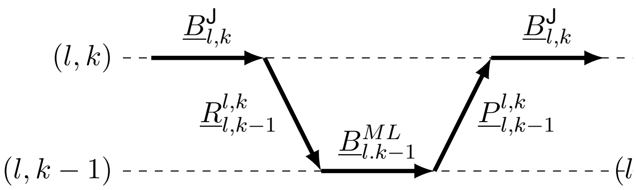


Figure 2. Two level in time coarsening strategy.

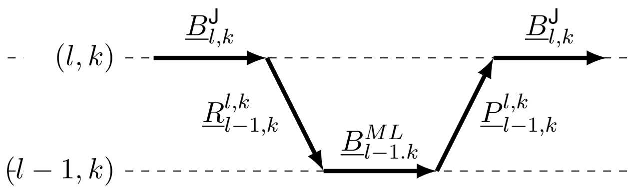


Figure 3. Two level in space coarsening strategy.

Algorithm 2 Multilevel preconditioner $\underline{c}_{l,k} = \underline{B}^{\mathsf{ML}}_{l,k}\underline{r}_{l,k}$ with smoother $\underline{B}^{\mathsf{SM}}_{l,k} = \underline{B}^{\mathsf{J}}_{l,k}$ or $\underline{B}^{\mathsf{GS}}_{l,k}$

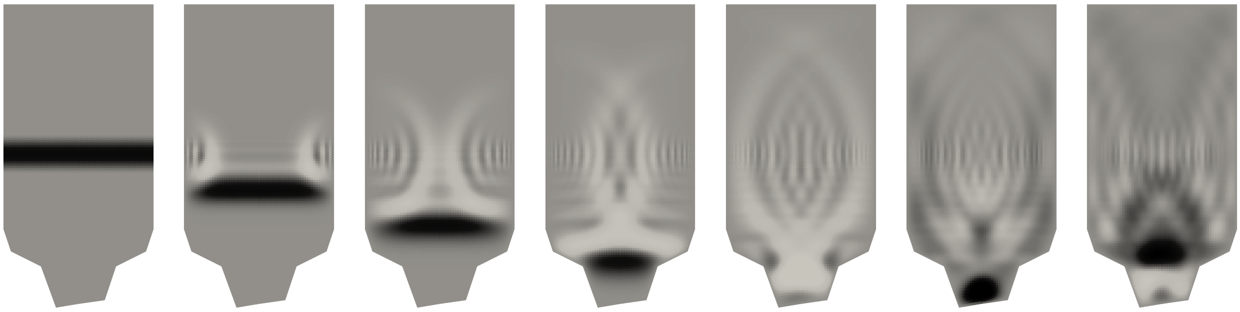
$\underline{c}_{l,k} = \underline{0}$
for $\nu = 1, \dots, \nu^{\mathsf{pre}}_{lk}$ **do**
  $\underline{w}_{l,k} = \underline{B}^{\mathsf{SM}}_{l,k}\underline{r}_{l,k}$
  $\underline{c}_{l,k} := \underline{c}_{l,k} + \underline{w}_{l,k}$ and $\underline{r}_{l,k} := \underline{r}_{l,k} - \underline{L}_{l,k}\underline{w}_{l,k}$
$\underline{r}_{l-1,k} = \underline{R}^{l,k}_{l-1,k}\underline{r}_{l,k}$ or $\underline{r}_{l,k-1} = \underline{R}^{l,k}_{l,k-1}\underline{r}_{l,k}$
$\underline{c}_{l-1,k} = \underline{B}^{\mathsf{ML}}_{l-1,k}\underline{r}_{l-1,k}$ or $\underline{c}_{l,k-1} = \underline{B}^{\mathsf{ML}}_{l,k-1}\underline{r}_{l,k-1}$
$\underline{w}_{l,k} = \underline{P}^{l,k}_{l-1,k}\underline{c}_{l-1,k}$ or $\underline{w}_{l,k} = \underline{P}^{l,k}_{l,k-1}\underline{c}_{l,k-1}$
$\underline{c}_{l,k} := \underline{c}_{l,k} + \underline{w}_{l,k}$ and $\underline{r}_{l,k} := \underline{r}_{l,k} - \underline{L}_{l,k}\underline{w}_{l,k}$
for $\nu = 1, \dots, \nu^{\mathsf{post}}_{lk}$ **do**
  $\underline{w}_{l,k} = \underline{B}^{\mathsf{SM}}_{l,k}\underline{r}_{l,k}$
  $\underline{c}_{l,k} := \underline{c}_{l,k} + \underline{w}_{l,k}$ and $\underline{r}_{l,k} := \underline{r}_{l,k} - \underline{L}_{l,k}\underline{w}_{l,k}$

7 A numerical experiment

The code is installed (with user m++ and password m++) and the examples are started by

```
svn co https://svn.math.kit.edu/svn/M++/SummerSchool
cd SummerSchool
make TimeStepping
make SpaceTime
mpirun -n 4 M++TimeStepping
mpirun -n 4 M++SpaceTime
paraview
```

The results are found in the log-files in the directory `log`, vtk-files are in `data/vtk` and can be viewed with `paraview`. Parameters of the configuration can be changed in `TimeStepping/conf/acoustic.conf` and `SpaceTime/conf/spacetime_acoustic.conf`.



$t = 0$ $t = 1$ $t = 2$ $t = 3$ $t = 4$ $t = 5$ $t = 6$

Figure 4. Solution at different time steps.

Problem configuration We consider an acoustic wave in $\Omega \subset (0,4) \times (-2.1, 6) \subset \mathbb{R}^2$ as given in Fig. 4. At $t = 0$ we start with the initial conditions

$$p_0(x_1, x_2) = \begin{cases} 100 \exp(-4(P_{\text{mid}} - x_2)^2)\,(1 - 4(P_{\text{mid}} - x_2)^2) & \text{if } 2|P_{\text{mid}} - x_2| < 1, \\ 0 & \text{else}, \end{cases}$$

$$\mathbf{v}_0(x_1, x_2) = \begin{cases} \Big(0, -100 \exp(-4(P_{\text{mid}} - x_2)^2)\,(1 - 4(P_{\text{mid}} - x_2)^2)\Big)^\top & \text{if } 2|P_{\text{mid}} - x_2| < 1, \\ \Big(0, 0\Big)^\top & \text{else} \end{cases}$$

for all $\mathbf{x} = (x_1, x_2)^\top \in \Omega$. The location in x_2-direction of the plane wave is controlled by the variable $P_{\text{mid}} \in \mathbb{R}$. The final time is $T = 6$, and the right-hand side $\mathbf{f} = \mathbf{0}$.

We consider the linear error functional in the region of interest $S = \{T\} \times (1,3) \times (-1,0)$

$$E(p, \mathbf{v}) = \frac{1}{|S|} \int_S p \,\mathrm{d}\mathbf{x}$$

Challenge Compute the value $E(p, \mathbf{v})$ up to an accuracy of approx. 1% with time stepping methods (see Appendix B) on uniform meshes and with the adaptive space-time method. Find out by numerical experiments which time step size and which polynomial degree is required for the time stepping method to achieve this accuracy. The same accuracy should by obtained with less degrees of freedom with the fully adaptive space-time method.

Appendix A: An identity for Legendre polynomials

Let $\lambda_{I,k} \in \mathbb{P}_k$ be the orthonormal Legendre polynomials with respect to the inner product in $\mathrm{L}_2(I)$ in the interval $I = (t_{n-1}, t_n)$.

Lemma 7.1 *We have* $\big(t\partial_t \lambda_{I,k}, \lambda_{I,k}\big)_{0,I} = k$ *for* $k \geq 0$*.*

Proof. We prove the result for the orthonormal Legendre polynomials $\lambda_k \in \mathbb{P}_k$ in $\mathrm{L}_2(-1,1)$; then, the general case follows directly from $\lambda_{I,k}(t) = \sqrt{\frac{2}{t_n - t_{n-1}}}\lambda_k\Big(2\frac{t-t_{n-1}}{t_n-t_{n-1}} - 1\Big)$.

Starting with $\lambda_{-1} \equiv 0$ and $\lambda_0 \equiv 1/\sqrt{2}$, we obtain recursively

$$(k-m+1)\partial_t^m \lambda_{k+1}(t) = (2k+1)t\partial_t^m \lambda_k(t) - (k+m)\partial_t^m \lambda_{k-1}^{(m)}(t)\,, \qquad k > 0\,,\ m \geq 0\,,$$

see [AS64, Lem. 8.5.3]. We have $\partial_t \lambda_0 \equiv 0$. For $k \geq 0$ we obtain from $(k+1)\lambda_{k+1}(t) = (2k+1)t\lambda_k(t) - k\lambda_{k-1}(t)$

$$(k+1)\partial_t \lambda_{k+1}(t) = (2k+1)\lambda_k(t) + (2k+1)t\partial_t \lambda_k(t) - k\partial_t \lambda_{k-1}(t)\,.$$

Subtracting $k\partial_t \lambda_{k+1}(t) = (2k+1)t\partial_t \lambda_k(t) - (k+1)\partial_t \lambda_{k-1}(t)$ results in $\partial_t \lambda_{k+1}(t) = (2k+1)\lambda_k(t) + \partial_t \lambda_{k-1}(t)$. This yields the assertion by

$$\big(t\partial_t \lambda_{k+1}, \lambda_{k+1}\big)_{0,(-1,1)} = \big(t(2k+1)\lambda_k, \lambda_{k+1}\big)_{0,(-1,1)} = \big((k+1)\lambda_{k+1}, \lambda_{k+1}\big)_{0,(-1,1)} = k+1\,.$$

$\square$

Appendix B: Time integration for linear systems

We consider time integration methods for the discrete evolution equation (7) in the form

$$\underline{M}\partial_t\underline{u} + \underline{A}\,\underline{u} = \underline{0}\,, \qquad \underline{u}(0) = \underline{u}_0\,, \tag{17}$$

where the (symmetric, positive definite, block-diagonal) mass matrix $\underline{M}$ and the non-symmetric stiffness matrix $\underline{A}$ with respect to a discontinuous finite element basis $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \dots$ are defined by

$$\underline{M} = \Big((M_h\boldsymbol{\phi}_k, \boldsymbol{\phi}_j)_{0,\Omega}\Big)_{j,k}\,, \qquad \underline{A} = \Big((A_h\boldsymbol{\phi}_k, \boldsymbol{\phi}_j)_{0,\Omega}\Big)_{j,k}\,.$$

The coefficient vector of the solution at time t with respect to the finite element basis and the corresponding element function $\mathbf{u}_h$ are denoted by

$$\underline{u}(t) = \big(u_j(t)\big)_j \in \mathbb{R}^{\dim H_h}\,, \qquad \mathbf{u}_h(t) = \sum\nolimits_j u_j(t)\boldsymbol{\phi}_j \in H_h\,.$$

The solution of this finite dimensional linear problem is given by

$$\underline{u}(t) = \exp(-t\underline{M}^{-1}\underline{A})\underline{u}_0\,, \quad t \geq 0\,, \tag{18}$$

where $\exp(\cdot)$ is the matrix exponential function. For a fixed time step $\tau > 0$ we compute approximations $\underline{u}^n \approx u(t_n)$ for $t_n = n\tau$. For one-step methods the approximations to the solution of (17) can be written as

$$\underline{u}^{n+1} = \Phi_n(-\tau\underline{M}^{-1}\underline{A})\underline{u}^n\,, \qquad n = 0, 1, \dots\,, \tag{19}$$

where Φ_n denotes the stability function of the method.

Explicit Runge-Kutta methods For an m-stage explicit Runge-Kutta method, Φ_n is a fixed polynomial of degree m, which approximates the exponential function in a neighborhood of zero. For instance, for the classical forth-order Runge-Kutta method we have $m = 4$ and

$$\Phi_n(\xi) = 1 + \xi + \frac{1}{2}\xi^2 + \frac{1}{6}\xi^3 + \frac{1}{24}\xi^4\,, \qquad \text{for all } n\,.$$

Each time step requires m multiplications with $\underline{A}$ and m solutions of linear systems with the block-diagonal matrix $\underline{M}$. These methods are simple to implement and computationally cheap, but the main disadvantage is the stability issue: all explicit Runge-Kutta schemes have a bounded stability region requiring time steps proportional to h^{-1} for first-order systems (CFL condition).

Implicit Runge-Kutta methods Implicit m-stage Runge-Kutta methods use a fixed rational function Φ_n with numerator and denominator degree at most m to approximate the exponential function. For hyperbolic problems as considered in this paper, Gauß collocation methods are particularly attractive [HW96, Chap. IV]. Here, Φ_n is the (m, m) Padé approximation to the exponential function. Gauß methods are A-stable and thus do not suffer from restrictions on the time step size τ for stability reasons. For $m = 1$ the the implicit midpoint rule is given by

$$\underline{u}^{n+1} = \underline{u}^n - \tau\big(\underline{M} + \tfrac{\tau}{2}\underline{A}\big)^{-1}\underline{A}\,\underline{u}^n\,.$$

For $m > 1$, a factorization into (in general complex) linear factors numerator and denominator is required. Each time step requires one matrix-vector multiplication with A, two with $\underline{M} + \gamma\tau\underline{A}$ for some (complex) coefficient γ and m solutions of linear systems with such coefficient matrices. Note that the stability property (12) for the upwind discretization shows that the linear system is dissipative, so that the implicit Gauss collocation methods are well-defined for all time steps.

Polynomial Krylov methods An alternative to explicit or implicit Runge-Kutta methods, for which the stability function Φ_n in (19) is fixed for all time steps, is to choose Φ_n adaptively. This can be accomplished by Krylov subspace methods. Standard Krylov subspace methods compute an approximation to $\underline{x} = \exp(-\tau \underline{M}^{-1}\underline{A})\underline{u}^n$ in the polynomial Krylov space

$$\mathcal{K}_m := \mathcal{K}_m(\underline{M}^{-1}\underline{A}, \underline{u}^n) = \operatorname{span}\left\{\underline{u}^n, \underline{M}^{-1}\underline{A}\,\underline{u}^n, \dots, (\underline{M}^{-1}\underline{A})^{m-1}\underline{u}^n\right\}.$$

The approximation proceeds in two steps. First, a basis of $\mathcal{K}_m$ is computed by the Lanczos or by the Arnoldi algorithm. Here, we only consider the Arnoldi algorithm with respect to the inner product $(\cdot,\cdot)_{\underline{M}}$. This yields a matrix $\underline{V}_m = [\underline{v}_1, \dots, \underline{v}_m] \in \mathbb{R}^{N\times m}$ and an upper Hessenberg matrix $H_m \in \mathbb{R}^{m\times m}$ such that

$$\underline{A}\,\underline{V}_m = \underline{M}\,\underline{V}_m\,\underline{H}_m + h_{m+1,m}\underline{M}\,\underline{v}_{m+1}\,\underline{e}_m^T, \quad \underline{V}_m^T\underline{M}\,\underline{V}_m = \underline{I}_m\,. \tag{20}$$

The $\underline{M}$-orthogonality of $\underline{V}_m$ shows that $\underline{H}_m = \underline{V}_m^T\underline{A}\,\underline{V}_m$. Now the approximation is given as

$$\exp(-\tau\underline{M}^{-1}\underline{A})\underline{u}^n \approx \underline{V}_m \exp(-\tau\underline{H}_m)\underline{V}_m^T\underline{M}\,\underline{u}^n\,,$$

see [GS92, Saa92]. Inserting $\underline{V}_m^T\underline{M}\,\underline{u}^n = \|\underline{u}^n\|_{\underline{M}}\,\underline{e}_1$ this yields the polynomial approximation

$$\underline{u}^{n+1} = \|\underline{u}^n\|_{\underline{M}}\,\underline{V}_m \exp(-\tau\underline{H}_m)\underline{e}_1 = \Phi_n(-\tau\underline{M}^{-1}\underline{A})\underline{u}^n \tag{21}$$

for some polynomial Φ_n of degree at most $m-1$, which is chosen automatically.

Algorithm 3 Polynomial Krylov method

1: **Input:** $\underline{M}$, $\underline{A}$, $\underline{v}$, τ, MaxIter, Tol
2: **Output:** $\underline{x}_m \approx \exp(-\tau\underline{M}^{-1}\underline{A})\underline{v}$, $m \le$ MaxIter, estimated error $\le$ Tol
3: $\beta = \|\underline{v}\|_{\underline{M}}\,, \quad \underline{v}_1 = \underline{v}/\beta$
4: **for** $m = 1, 2, \dots,$ MaxIter **do**
5: $\quad \underline{w} = \underline{A}\,\underline{v}_m$
6: $\quad$ solve $\underline{M}\,\underline{v}_{m+1} = \underline{w}$
7: $\quad$ **for** $k = 1, \dots, m$ **do**
8: $\quad\quad h_{k,m} = \underline{v}_k^T\underline{w}$
9: $\quad\quad \underline{v}_{m+1} = \underline{v}_{m+1} - h_{k,m}\underline{v}_k$
10: $\quad h_{m+1,m} = \left\|\underline{v}_{m+1}\right\|_{\underline{M}}$
11: $\quad \underline{v}_{m+1} = \underline{v}_{m+1}/h_{m+1,m}$
12: $\quad \underline{y}_m = \beta\exp(-\tau\underline{H}_m)\underline{e}_1$
13: $\quad \delta_m = \left\|\underline{y}_m - [\underline{y}_{m-1}; 0]\right\| / \left\|\underline{y}_m\right\|$
14: $\quad \epsilon_m = 1 + \left\|\underline{y}_m\right\|$
15: $\quad$ **if** $\delta_m < 1$ **then**
16: $\quad\quad \epsilon_m = \min\left(1 + \left\|\underline{y}_m\right\|, \delta_m/(1-\delta_m)\left\|\underline{y}_m\right\|\right)$
17: $\quad$ **if** $\epsilon_m \le$ Tol **then**
18: $\quad\quad$ break
19: **if** $m \ge$ MaxIter and $\epsilon_m >$ Tol **then**
20: $\quad$ no convergence
21: $\underline{x}_m = [\underline{v}_1, \dots, \underline{v}_m]\underline{y}_m$

The stopping criteria in Line 17 of Alg. 3, was introduced in [vdEH06], see also [BGH13] for a detailed investigation of residuals of the matrix exponential. Here, δ_m is an estimation of the relative error $\|\underline{x}_m - \underline{x}\|_{\underline{M}} / \|\underline{x}_0 - \underline{x}\|_{\underline{M}}$ in the mth Krylov step. Note that $\underline{y}_m$ has to be measured in the Euclidean norm; since for $\underline{x}_m = \underline{V}_m\underline{y}_m$, we have $\|\underline{x}_m\|_{\underline{M}} = \|\underline{y}_m\|$.

References

[AS64] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Applied mathematics series. Dover Publications, 1964.

[BGH13] M. A. Botchev, V. Grimm, and M. Hochbruck. Residual, Restarting, and Richardson Iteration for the Matrix Exponential. *SIAM J. Sci. Comput.*, 35(3):A1376–A1397, 2013.

[BR99] W. Bangerth and R. Rannacher. Finite element approximation of the acoustic wave equation: Error control and mesh adaptation. *East-West J. Numer. Math.*, 7(4):263–282, 1999.

[BR03] W. Bangerth and R. Rannacher. *Adaptive finite element methods for differential equations*. Birkhäuser Verlag, Basel, 2003.

[Bra07] D. Braess. *Finite Elements. Theory, fast solvers, and applications in solid mechanics. 3th ed.* Cambridge University Press, Cambridge, 2007.

[DFW16] Willy Dörfler, Stefan Findeisen, and Christian Wieners. Space-time discontinuous galerkin discretizations for linear first-order hyperbolic evolution systems. *Comput. Methods Appl. Math.*, 16(3):409–428, 2016.

[EN06] Klaus-Jochen Engel and Rainer Nagel. *A short course on operator semigroups*. Universitext. Springer, New York, 2006.

[GS92] E. Gallopoulos and Y. Saad. Efficient solution of parabolic equations by Krylov approximation methods. *SIAM J. Sci. Statist. Comput.*, 13(5):1236–1264, 1992.

[HPS$^+$15] M. Hochbruck, T. Pazur, A. Schulz, E. Thawinan, and C. Wieners. Efficient time integration for discontinuous Galerkin approximations of linear wave equations. *ZAMM*, 95(3):237–259, 2015.

[HR03] V. Heuveline and R. Rannacher. Duality-based adaptivity in the hp-finite element method. *J. Numer. Math.*, 11:95–113, 2003.

[HW96] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*, volume 14 of *Springer Series in Computational Mathematics*. Springer, Berlin, Heidelberg, 2nd edition, 1996.

[HW08] J. S. Hesthaven and T. Warburton. *Nodal Discontinuous Galerkin Methods*. Springer, 2008.

[Lev02] Randall J. Leveque. *Finite volume methods for hyperbolic problems.* Cambridge: Cambridge University Press, 2002.

[RR04] Michael Renardy and Robert C. Rogers. *An introduction to partial differential equations. 2nd ed.* New York, NY: Springer, 2004.

[Saa92] Y. Saad. Analysis of some Krylov subspace approximations to the matrix exponential operator. *SIAM J. Numer. Anal.*, 29(1):209–228, 1992.

[vdEH06] J. van den Eshof and M. Hochbruck. Preconditioning Lanczos approximations to the matrix exponential. *SIAM J. Sci. Comput.*, 27(4):1438–1457, 2006.